

# WEST Search History

DATE: Tuesday, December 31, 2002

<u>Set Name</u>	<u>Query</u>	<u>Hit Count</u>	<u>Set Name</u>
side by side			result set
<i>DB=USPT,PGPB,EPAB,DWPI,TDBD; THES=ASSIGNEE; PLUR=YES; OP=ADJ</i>			
L2	L1 and carrot	16	L2
L1	antifreeze adj (polypeptide? or peptide? or protein?)	98	L1

END OF SEARCH HISTORY

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2002, 18:56:19 : Search time 1305.46 Seconds

(without alignments)  
12624.425 Million cell updates/sec

Title: US-09-308-140-6

Sequence: 1 atgaattatgaatcattt.....ctccatgccgaatgctag 999

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapept 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_sy:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htgo\_hum:\*  
31: em\_htgo\_inv:\*  
32: em\_htgo\_hum:\*  
33: em\_htg\_hum:\*  
34: em\_htg\_inv:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	999	100.0	999	6	A91926
2	999	100.0	1140	8	AF055480
3	999	100.0	1238	8	AF055480
4	584.4	58.5	829	6	DCA131340
5	356.2	35.7	1285	8	AF19131
6	332	33.2	1162	8	ADPGIP
7	328.8	32.9	1057	8	MDU77041
8	328.8	32.9	1057	8	PPRPG1
9	295.4	29.6	1058	6	128276
10	290.4	29.1	1002	8	AF305093
11	290.4	29.1	2075	6	128277
12	284	28.4	909	8	TOMPIR
13	284	28.4	909	8	AF159167
14	284	28.4	909	8	AF159170
15	284	28.4	909	8	AF159171
16	283.2	28.3	1207	8	CSPGIP
17	283.2	28.3	984	8	AB015643
18	283.2	28.3	984	8	AB016204
19	281.6	28.2	984	8	AB016206
20	280.8	28.1	909	8	AB015198
21	280.2	28.0	988	8	AF159169
22	280	28.0	988	8	AB013397
23	280	28.0	984	8	AB015356
24	278.4	27.9	984	8	AB016205
25	277.6	27.8	909	8	AB020529
26	277.6	27.7	988	8	AF159168
27	258.4	25.9	1034	8	AY048295
28	256.8	25.7	993	8	AF229249
29	249.2	24.9	981	8	AF229250
30	249.2	24.9	1159	8	AY035121
31	245.2	24.5	914	8	AB021791
32	245	24.5	751	8	AF196906
33	241.8	24.2	751	8	AF196879
34	240.6	24.1	751	8	AF196884
35	240.2	24.0	751	8	AF196880
36	240.2	24.0	751	8	AF196881
37	240.2	24.0	751	8	AF196890
38	238.6	23.9	751	8	AF196929
39	238.6	23.9	751	8	AF196932
40	238.2	23.8	751	8	AF196907
41	237	23.7	751	8	AF196901
42	237	23.7	751	8	AF196902
43	237	23.7	751	8	AF196931
44	235.6	23.7	751	8	AF196867
45	235.4	23.6	751	8	AF196900

## ALIGNMENTS

RESULT 1  
A91926  
LOCUS A91926 999 bp DNA  
DEFINITION Sequence 6 from Patent WO9822591.  
ACCESSION A91926  
VERSION A91926.1 GI:6740789  
KEYWORDS  
SOURCE  
ORGANISM  
Daucus carota

PAT 22-JAN-2000

REFERENCE  
AUTHORS Byass, L.J. and Doucet, C.J.  
TITLE CARROT ANTIFERREZE POLYPEPTIDES  
JOURNAL Patent: WO.9822591-A 6 28-MAY-1998.  
BYASS LOUISE JANE (GB); DOUCET CHARLOTTE JULIETTE (GB)

FEATURES  
source  
1..999  
/organism="Daucus carota"

CDs  
/strain="AUTUMN KING"  
/db\_xref="taxon:4039"  
1. 999  
/codon\_start=1  
/number=1  
/evidence="experimental"  
/product="ANTI FREEZE PROTEIN"  
/protein\_id="CAB69452.1"  
/db\_xref="GI:6740790"

/translation="MNIESFCPIICMIFLCIPNLISQRCNNNDKALLOIKTAL  
KNPTITDSWVSDDCGMDLVECDENSRNIIISLIQDDALTGQIPQVGDLPYLOAL  
WFKRLPNLFGKIPEIISALDKLSRLSTSLSGVPLFPOLTKITCLDLSFNKILG  
VIPOLSTLPNKLALHLEBNELTGELPIDIFGNFAGSPDIYLSHNOITGVFKTFARAD  
PIRDFSGNRLEGDISFLFGPKRLEMDLPSGNVLSFNSVOETFPSTLYLDLNHO  
ISSLSSESLAKLIDTQFNVSQNNLCGKIPTGQNLQRFDRYALVHNSCLGAPLPEC"

BASE COUNT 277 a 266 c 185 g 271 t  
ORIGIN

Query Match 100.0% Score 999: DB 6: Length 999:

Best Local Similarity 100.0%: Pred. No. 7.3e-287: Mismatches 0: Indels 0: Gaps 0:

Matches 999: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 atgaataatgaatcaatcttctgcccattttgtgcatatgcatatcttccctgccc  
1 ATGAATATGAATCATCTCTCTGCCCCATTTTGTGATGATGATTTCTCTGCCCC  
60  
QY 61 ccaaacctctctgcatcacaaagaatgacaagaagaagcattaccccaatc 120  
61 CCAAACTCTCTCATCACAAAGATGCAACACGACACAGCAAGCTTTACTCCAAATC  
120  
QY 121 aaaaagccttgaaaaaaccaccatcacagcatcagtgctcagaagcagatctgtgt 180  
121 AAAAGAGCCTTGAAAAACCCACCATTACAGACTCATGGGTGTGACAGACGATTTGTGT  
180  
QY 121 aaaaagccttgaaaaaaccaccatcacagcatcagtgctcagaagcagatctgtgt 180  
121 AAAAGAGCCTTGAAAAACCCACCATTACAGACTCATGGGTGTGACAGACGATTTGTGT  
180  
QY 181 ggttgagacatgctgaatgtagaagaacacagcaacgcataatctccatcaatcaa 240  
181 GGTGTGACCATCTGATCGAATGTGAGCAACACGACGCAATTAATTTCCCTCATTAATTCAA  
240  
QY 241 gaagcagaagaactctcaccggcacaatccaccctcaggttgagagaacacatccctcaa 300  
241 GACGACCAACCTCTCACCGCCCAAAATCCACGTACGTGGGAAACCTCACCATACCTCCAA  
300  
QY 301 gacctatgcttcgtaaacctcccaatcttctcgaaaaatccccagaagaactcttgcga 360  
301 GCCTTATGTTCCGTTAACTCCCAATCTTTTGGAAAAATCCACAGAGAAATTTCTGCA  
360  
QY 361 ctcaagaacctaataatccctcagaactcagcagcagctcagtgccctgtcccttta 420  
361 CTCAAAGACCTTAATAATCCCTCAGACACTGCTCAGACAGTCTCAGTGGCCCTGTCCCTTA  
420  
QY 421 tttctccctgaagcttgcgaagaacttgttttagactatcgtttacaaactttggggt 480  
421 TTTCTTCCCTGAGCTTACGAACACTTGTGTTAGACTTATCGTTTAAACAACTTTTGGGT  
480  
QY 481 gtaatccctcctcagcttctcactcttccgaaccttaaaagccctgcacttagaagctaac 540  
481 GTATATCCCTCTCAGCTTTCGACTCTTCGACACCTTAACCCCTGCACCTTAGAAGCTAAC  
540  
QY 541 gaaactcagtgtaaatcccgatattcttgggaatttgcgtgagatcccgagacataat 600  
541 GAACCTCACCGGTGAAATCCCGATATCTTTGGGAATTTTCTGTGATCCCGGACATATAT  
600  
QY 601 ctctcgatatcagcagctcagcgggttgcctccaaaactttgcctgagagagatccaatt 660  
601 CTTTCGATATACAGCTCACCGGGTGTGTTCCAAACTTTTCTGTAGAGAGATTCACAAAT  
660  
QY 661 aagctcagcttctcaggaagaacagagatgatatcatcttctgtttgggctaaa 720  
661 AGGCTGACATTTCTCAGGAGACACAGATGATATTTCTTGTGGGCTTAA  
720  
QY 721 aaacgcttggaatgtagatttttcaaggaacgctgcttagttcaattctccaggggtg 780  
721 aaacgcttggaatgtagatttttcaaggaacgctgcttagttcaattctccaggggtg  
780

DB 721 AAACGCTGGAATCTAGATTTTTCAGGAACGCTCTAGTTCAATTTCTCCAGGGTG 780  
QY 781 caggaattccaccctcttgacatacttagacttgaaccataaccagatcagcgaagt 840  
781 CAGGAGTTTCCACCTCTTTGACATCTTAGATTTGACCATTAACCATACCGGAACT 840  
QY 841 ctctcagtagaatttgctaaattgagacctgcagacatttaacgtcagatataatctc 900  
841 CTCTCAGTAGAATTTGGCTTAATTTGACACCTGCAGACATTAACGTGATATATATCTC 900  
QY 901 tgcggcagaattcccaacagggggaacctccagagatcgcagcgtacggcctatccac 960  
901 TCGGCGCAAAATTTCCAAACAGGGGGAACCTCCAGAGATTGACCGCTACGCTATCTCCAC 960  
QY 961 aacagcttgctgtgtgtgtctccatgcccagaatgctag 999  
961 AACAGTTGCTGTGTGTCTCTCCATTGCCCAGAAATGCTAG 999

RESULT 2

AF055480 1140 bp mRNA PLN 06-OCT-1998

DEFINITION Daucus carota antifreeze protein (AFP) mRNA, complete cds.

ACCESSION AF055480

VERSION AF055480.1 GI:3702802

KEYWORDS

SOURCE carrot.

ORGANISM Daucus carota

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids II; Apiales; Apiaceae; Daucus.

(bases 1 to 1140)

Worrall,D., Elias,L., Ashford,D., Smallwood,M., Sidebottom,C.,

Hillford,P., Telford,J., Holt,C. and Bowles,D.

A carrot leucine-rich-repeat protein that inhibits ice

recrystallization

Science 282 (5386), 115-117 (1998)

98429644

2 (bases 1 to 1140)

Worrall,D.

Direct Submission

Submitted (23-MAR-1998) Biology, University of York, University

Road, PO Box 373, York YO1 5YW, UK

Location/Qualifiers

1. 1140

/organism="Daucus carota"

/cultivar="Autumn King"

/db\_xref="taxon:4039"

/tissue\_type="tap root"

/note="cold acclimated"

1. 1140

/gene="AFP"

13. 1011

/gene="AFP"

/note="similar to plant polygalacturonase inhibitor

protein; PCIP"

/codon\_start=1

/product="antifreeze protein"

/protein\_id="AAC62932.1"

/db\_xref="GI:3702803"

/translation="MNIESFCPIICMIFLCIPNLISQRCNNNDKALLOIKTAL

KNPTITDSWVSDDCGMDLVECDENSRNIIISLIQDDALTGQIPQVGDLPYLOAL

WFKRLPNLFGKIPEIISALDKLSRLSTSLSGVPLFPOLTKITCLDLSFNKILG

VIPOLSTLPNKLALHLEBNELTGELPIDIFGNFAGSPDIYLSHNOITGVFKTFARAD

PIRDFSGNRLEGDISFLFGPKRLEMDLPSGNVLSFNSVOETFPSTLYLDLNHO

ISSLSSESLAKLIDTQFNVSQNNLCGKIPTGQNLQRFDRYALVHNSCLGAPLPEC"

BASE COUNT 332 a 284 c 208 g 316 t

ORIGIN

Query Match 100.0% Score 999: DB 8: Length 1140:  
Best Local Similarity 100.0%: Pred. No. 7.4e-287: Mismatches 0: Indels 0: Gaps 0:  
Matches 999: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

```

QY 1 atgaatatgaatcatcttctgacctatatttggcataagcatgatttctctgacct 60
Db 13 ATGAATATTGAATATCTTTCTGCGCTATTTTGGCATATGACATGATTTCTCTGCTT 72
QY 61 ccaaacctctctgcatcaaaaagatgcaacaacagacagagaagatttactccaatc 120
Db 73 CCAAACTCTCTGCATGCACAAAGATGCACAAACAGACAGACAGCTTATCTCAATC 132
QY 121 aaaaagccttgaaaaaaccaccatatacagactatgggtgcagacagatgtttt 180
Db 133 AAAACAGCCTTGAAAAACCCACATTAAGACTCATGGGTGCACAGACAGCTTGTG 192
QY 181 ggttggaactaagtgaatgtagcgaacacagacacacacacacacacacacacac 240
Db 193 GGTGGGACCTTAAGTGCAGAAACAGACACACACACACACACACACACACACACAC 252
QY 241 gacgacgaagctctcagcagcaaatcccaactcagtgtaggagactatacactc 300
Db 253 GACGACGAGAGCTCTCAGCGGCAATCCACCTCAGTGAGGAGACCTACCAATCTCA 312
QY 301 gacctatggttcgcaaacctcccaactcttctcggaaaaaatcccgagaatctt 360
Db 313 GCGTTATGTTCCGTAAACTCCCAATCTTTCCGAAAAAATCCAGAAATTTCTGCA 372
QY 361 ccaagaagactaaatccctcagactcagctcagcagctcagctcagctcagctc 420
Db 373 CTCAGAGACTTAATATCCCTCAGACTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 432
QY 421 tctctccctcagactcagaaactaacttggtttagaactatcgtttaaaaaact 480
Db 433 TCTTCCCTCAGCTTACCAAACTAATCTTTAGACTTATGTTTAAACAACTTTGGGT 492
QY 481 gtaacccctcagctcagcttccactcttcagacacttaaacgccgcaacttaga 540
Db 493 GTAATCCCTCTCAGCTTTCACACTTCCGAACTTAAGCCCTGCACCTTAGAACG 552
QY 541 gaactacagcgttagaattcccgatatactcttggaatttgctgagatcccgaca 600
Db 553 GAACCTACCGGTGAATCCCGATATCTTGGGAATTTCTGGATCCCGACATATAT 612
QY 601 ctttcgataacagctcagcaggttgctcccaaaacttctcagagacagatcca 660
Db 613 CTTTCGATTAACAGCTCAGCGGCTTGTTCACAAACTTTTCTAGAGAGATCCAA 672
QY 661 agctcagactctcaggaacagactagaagtgcatttcaactcttctgttgag 720
Db 673 AGGCTTGACTTTCAGAGGAACAGACTAGAGATGATTTCTTGTGCGCTTAA 732
QY 721 aaagccttgaaatgtagatttttcaagaaaagtgtagttcaattctccaggtg 780
Db 733 AAACGCTTGAAATGCTATATTTTCAGGAAACGTCCTTAGTTCAATTTCCAGGG 792
QY 781 caggaagtttccaccctcttgacatattagacttgaaacataacacagacag 840
Db 793 CAGGAGTTTCCACCTCTTTGACATTAAGCTTGAACCATACAGATCAGCGAAGT 852
QY 841 ctgtcagatgaattggtcaaatatgacactgcagacatttaagtgatgataata 900
Db 853 CTGTGAGATGATTTGGTAATGTGACCTGCAGACATTTAACGTCAGATTAATAT 912
QY 901 tgcgcaagatctcaaacgggggaaacactcagagatcgacgtagcgtatctcac 960
Db 913 TCGGCAAGATTTCAAAAGGGGGAACCTCCAGAGATTCGACCGGAGGCTATCTCAC 972
QY 961 aacagttgcttggtggtcgaattgccaagaatgctag 999
Db 973 AACAGTTGCTTGTGTGCTCCTCAATTCGCAAGATGCTAG 1011

```

RESULT 3  
DCA131340 1238 bp DNA PLN 06-APR-1999  
LOCUS

```

DEFINITION Daucus carota afp gene.
ACCESSION AJ131340
VERSION AJ131340.1 GI:4455919
KEYWORDS afp gene; antifreeze polypeptide; leucine-rich repeat protein;
SOURCE Daucus carota
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
1 (bases 1 to 1238)
Meyer K., Kell M. and Nalder J.
A leucine-rich repeat protein of carrot that exhibits antifreeze
activity
FEBS Lett. 447 (2-3), 171-178 (1999)
99229753
REFERENCE 2 (bases 1 to 1238)
AUTHORS Meyer K.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-1998) Meyer K., Forestry Research Unit, Shell
International Renewables, HRI East Malling, Kent ME19 6 BJ, UK
FEATURES
source
1..1238
/organism="Daucus carota"
/cultivar="Newmarket F1"
/db_xref="taxon:4039"
1..31
/gene="afp"
1..1238
/gene="afp"
32..1030
/gene="afp"
/note="polylacturonase inhibitor-like protein;
leucine-rich repeat protein"
/region="start-1"
/product="antifreeze polypeptide"
/protein_id="CAB37347.1"
/db_xref="GI:4455920"
/translation="MNIESFPCICIMFICLRLNSASORCNNDKQALLQITL
KPTITDSVSDDCGMDIVECEDTSNRIISLIIDDEALTGQIPVQGLPYLQAL
WRRKIPNFGKIPETISALKDLSRLSTSLSGVPLFFQTLTGLDLSFNKLG
VIRPOLSTNPKALHLELNELTGEIPDIFGNFASPDVLSHNOITGTVKTFARAD
PIRDFSGNRLEKIDISFLEGPRLKLEMDFGSVLSFNSVQEPPEPLTYLDLNHO
ISGSLSELAKIDLOTFNVDNMLCGKIPITGNSLORFRTATVHNSCLCGAPLPEC"
32..109
/gene="afp"
110..1027
/gene="afp"
/product="polylacturonase inhibitor-like protein;
leucine-rich repeat protein"
1031..1238
/gene="afp"
BASE COUNT 359 a 218 g 361 t
ORIGIN
Query Match 100.0%; Score 999; DB 8; Length 1238;
Best Local Similarity 100.0%; Pred. No. 7,4e-287;
Matches 999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

Db 212 GGTGGGACCTAGTCGATGAGCAAAACGCAACCCCAATATTTCCCTCATATTCAA 271
QY 241 GAGCAGGAGCTCAGCAGGCAAAATCCACCTGAGTGGGAGACCTACCACTACCA 300
Db 272 GACGACGAGCTCAGCAGGCAAAATCCACCTGAGTGGGAGACCTACCACTACCA 331
QY 301 GECTATGATCGTAAATCCCAATCTCTCGGAAATCCCAAGAAATTTCTGCA 360
Db 332 GCTTATGATCGTAAATCCCAATCTCTCGGAAATCCCAAGAAATTTCTGCA 391
QY 361 CTCAAGAGCTAAATCCCTGAGACTCAGTCGACGACGATCGGCGCTGCTTTA 420
Db 392 CTCAAGAGCTAAATCCCTGAGACTCAGTCGACGACGATCGGCGCTGCTTTA 451
QY 421 TCTCCCTCAGCTAAGAACTAAGTCTGATGATGATGATGATGATGATGATGAT 480
Db 452 TCTCCCTCAGCTAAGAACTAAGTCTGATGATGATGATGATGATGATGATGAT 511
QY 481 GTAACTCCCTCAGCTAAGAACTAAGTCTGATGATGATGATGATGATGATGAT 540
Db 512 GTAACTCCCTCAGCTAAGAACTAAGTCTGATGATGATGATGATGATGATGAT 571
QY 541 GAATCAGCGGTGAATCCCGATATCTTGGGAATTTCTGATGATGATGATGAT 600
Db 572 GAATCAGCGGTGAATCCCGATATCTTGGGAATTTCTGATGATGATGATGAT 631
QY 601 CTCTCGATAACAGCTCAGCGGTGTTCTCCCAAACTTTGCTAGACGATCAAT 660
Db 632 CTCTCGATAACAGCTCAGCGGTGTTCTCCCAAACTTTGCTAGACGATCAAT 691
QY 661 AGGCTCAGCTCAGCGGTGAAGAGTGAATTTGATTTGATTTGATTTGATTT 720
Db 692 AGGCTCAGCTCAGCGGTGAAGAGTGAATTTGATTTGATTTGATTTGATTT 751
QY 721 AAGCGCTGGAATGCTAGATTTTCAAGAAAGTGTGATTTGATTTGATTTGAT 780
Db 752 AAGCGCTGGAATGCTAGATTTTCAAGAAAGTGTGATTTGATTTGATTTGAT 811
QY 781 CAAGAGTTCACGCTCTTGACATACTGACATACTGACATACTGACATACTGAC 840
Db 812 CAAGAGTTCACGCTCTTGACATACTGACATACTGACATACTGACATACTGAC 871
QY 841 CTCTCAGTGAATGCTGAATGCTGAATGCTGAATGCTGAATGCTGAATGCT 900
Db 872 CTCTCAGTGAATGCTGAATGCTGAATGCTGAATGCTGAATGCTGAATGCT 931
QY 901 TGGGCAAGATTCACAGGAGGGAACCTCCAGAGATTGACCGCTATCTCCAC 960
Db 932 TGGGCAAGATTCACAGGAGGGAACCTCCAGAGATTGACCGCTATCTCCAC 991
QY 961 AACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 999
Db 992 AACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1030

```

RESULT 4  
 LOCUS A91931 829 bp DNA  
 DEFINITION Sequence 11 from Patent WO9822591.  
 ACCESSION A91931  
 VERSION A91931.1 GI:6740793  
 KEYWORDS  
 SOURCE  
 ORGANISM *Daucus carota*  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids II; Apiales; Apiaceae; Daucus.  
 BYASS, L.J. and Doucet, C.J.  
 TITLE CARROT ANTIFREEZE POLYPEPTIDES  
 JOURNAL Patent: WO 9822591-A-11 28-MAY-1998;  
 BYASS LOUISE JANE (GB); DOUCET CHARLOTTE JULIETTE (GB)

```

FEATURES
    source          1..829
                    /organism="Daucus carota"
                    /strain="AUTUMN KING"
                    /db_xref="taxon:4039"
                    <1..594
    CDS
                    /codon_start=1
                    /product="ANTI FREEZE PROTEIN"
                    /protein_id="CAB69453.1"
                    /db_xref="GI:6740794"
                    /translation="GVPLPFPOLTKTICLDSFNKLLGVIPPOLSTLPNLKALHLER
                    NEITGEIPDIFGNFAGSPDILSHNLTGEVETKPARADPILDSGNLEDDISFLF
                    GKRRLEMDFGSNVLSFNSRVSQEPSTLTLDLNHNHNISSISSELAKLDELDTFNV
                    SNNICGKIPITGNGIQRFRRTAYLNHNSCLGAPLEPC"
BASE COUNT      235 a      176 c      160 g      258 t
ORIGIN
Query Match          58.5%; Score 584.4; DB 6; Length 829;
Best Local Similarity 99.0%; Pred. No. 3,4e-163;
Matches 588; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 406 ggcctgtcccttattcttccctcagcttgcgaacaaactagttttagactatcgltt 465
Db 1 GGGCGGTGGCGCTGTTCTTCCCTCAGCTTACGAACCTAAGTCTTATGCTTT 60
QY 466 aaaaacttttgggttlaacccctcctcagcttccacttccgaacttaagccctg 525
Db 61 AACAACTTTTGGGTATATCCCTCTCAGCTTCCAGCTTCCGAACCTTAAAGCCCTG 120
QY 526 caattagaacglaacgaactcaacggttgaatcccgatattcttgggaatttggctga 585
Db 121 CACTTGAACGTAACGACTCAGCGGTGAATCCCGATATCTTTGGAAATTTGCTGGA 180
QY 586 tcccggacatalatcttcgcataaacacgctcaccggttggltcccaaaactttgct 645
Db 181 TCCCGGACATATATCTTCCGATATACAGCTCAGCGGTTTGTTCCCAAACTTTGGCT 240
QY 646 agagcagatccaattaggtgctgaacttctcagggaaagactgaagtgatattcattc 705
Db 241 ACAGGATGCCAATTAAGCTGCACTTCTCAGGAAACAGACTGAAGTGATATTTTCATTC 300
QY 706 ttgtttggcccaaaaaagccttgggaatgctagatttttcaggaaagctgtttagtttc 765
Db 301 TTGTTGGGCTTAAAAAAGCTTGGAAATGCTAGATTTTTCAGAAAGCTGTTAGTTTC 360
QY 766 aattctccagggtgcaggagtttccaccctcttgcataacttgaacttgaaccataac 825
Db 361 AATTCTCCAGGAGTCCAGAGTTTCCACCCTCTTGGACATATTGACTTGAACCATTAAC 420
QY 826 cagatcagcgggaagctcgtgcagtgaaattgctaaattggacctgcagacattaaagtc 885
Db 421 CAGATCAGCGGAAGTCTGCGAGTAATTTGGCTAATTTGACACTTCAGACATTTAACTGC 480
QY 886 agtgaataatctcttgcggaagaattccaagaagggaacactccaagaattgacagct 945
Db 481 ACTGATTAATATCTTGGCGCAAGATTCACAGGGAACCTCCAGAGATTGACCGCT 540
QY 946 acggcctatctccacaacagtgctgtgtgtgtgtccatctgcagaaatgtgtag 999
Db 541 ACGGCTATCTCCACAACAGTTGCTGTGTGTGTGCTCATTTGCCAAGATGCTAG 594

```

RESULT 5  
 LOCUS ADGP1P 1285 bp mRNA  
 DEFINITION A.deliciosa pgjp mRNA for polygalacturonase inhibitor.  
 ACCESSION Z49063  
 VERSION Z49063.1 GI:1143380  
 KEYWORDS polygalacturonase inhibitor.  
 SOURCE Actinidia deliciosa.  
 ORGANISM *Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;*



TITLE Direct Submission  
JOURNAL Submitted (01-NOV-1996) USDA-ARS, Horticultural Crops Quality Laboratory, BARC-WEST, BLDG 002, Beltsville, MD 20705, USA

FEATURES  
Location/Qualifiers

source 1..1162  
/organism="Malus x domestica"

gene /db\_xref="taxon:3750"

CDS 1..1162  
/gene="mdg1p1"

19..1011  
/gene="mdg1p1"

/note="pg inhibitor"

/codon\_start=1

/product="polygalacturonase-inhibiting protein"

/protein\_id="AAB19212.1"

/db\_xref="GI:1679733"

/translation="MELKFSIFLSLTLFSSVLKPAISDLICNPDDKRVLLQIKKAFG  
PYVLASWKSDTDCDWCVTGDTNTRINSITIFAGOVSGOIPALVGDLPYLETFEH  
KQNLTPGIPQIAIKLKLKFLRSLMTLSSVDFLSQNLKTLFLDSFNLTGAIP  
SSLSOLPMLNALHLDNRKLTHIPKSLQFIGNVDPDLISHNLSGNIPISFAOMDFT  
SIDLSRNKLEGDASVIFGLNKTTQIYDLSRNLEFNLKVEFPTSLTSDINHNKITYG  
SIPVEFTLNQFLNVSYNRLCGQIPVGKIQSEDEYSYFHNRLCGAPLPSCK"

BASE COUNT 319 a 357 c 210 g 276 t

ORIGIN

Query Match 33.2% Score 332; DB 8; Length 1162;  
Best Local Similarity 60.4%; Pred. No. 7.1e-88;  
Matches 602; Conservative 0; Mismatches 385; Indels 9; Gaps 3;

QY 7 atgatacatcttctgacctatttggatagatgattctctcgcttccaac 66  
DB 19 atggaactcactgattctccctctccttacccttctcctcctcaaaa 78  
QY 67 cctctgcatcaagaagacacaaacagaagcagcttaccatcaaaa 126  
DB 79 cccgctctctccgattctgcacaccccgacacacacacacacacacac 138  
QY 127 gctctgaacaaacccacacacacacacacacacacacacacacacac 186  
DB 139 gccctgcggcgcacaccttggacacacacacacacacacacacacacac 198  
QY 187 gacctgactgattggaagacacacacacacacacacacacacacacac 246  
DB 199 tactgcgtcactctgacacacacacacacacacacacacacacacac 258  
QY 247 gaagctcagcggcgaacacacacacacacacacacacacacacacac 306  
DB 259 cagg--tatccggcccaatccacacacacacacacacacacacacac 315  
QY 307 tggctcgaac 366  
DB 316 gaattccac 375  
QY 367 gacctaaac 426  
DB 376 gcactcgaacttctcagcagcagcagcagcagcagcagcagcagcagc 435  
QY 427 cctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 486  
DB 436 agccac 495  
QY 487 cctcctcagcagcagcagcagcagcagcagcagcagcagcagcagc 546  
DB 496 cccagctcgcttctcagcagcagcagcagcagcagcagcagcagcagc 555  
QY 547 accggtgaaatcccgatattcttgggaatttgcgatcc---cggacata 603  
DB 556 acaggtcattatcccaaatccggttgaggacacacacacacacacacac 615  
QY 604 tcgacataac 663  
DB 616 tccac 675

QY 664 ctgacatctcagggagacagactagaagtgatattcattctgttggcctaaaaa 723  
DB 676 atagattatccaggaacacacacacacacacacacacacacacacacac 735  
QY 724 cggctggaatctgactgatttccaggaacgagcagtgattccaaattccaggg 783  
DB 736 ac 792  
QY 744 gatttccac 843  
DB 753 gacttccgac 852  
QY 844 tcgagtgatgagcagcagcagcagcagcagcagcagcagcagcagcagc 903  
DB 853 ccagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 912  
QY 904 ggcagatctcagcagcagcagcagcagcagcagcagcagcagcagcagc 963  
DB 913 gctcagattccagcagcagcagcagcagcagcagcagcagcagcagcagc 972  
QY 964 agtgctgt 999  
DB 973 ccagctgt 1008

RESULT 7  
PYPGI 1057 bp mRNA PLN 20-JUN-1994  
LOCUS Pyrus communis polygalacturonase inhibitor mRNA, complete cds.  
ACCESSION L09264  
VERSION L09264.1 GI:169683  
KEYWORDS polygalacturonase-inhibiting protein.  
SOURCE Pyrus communis (strain Bartlett) immature fruit fruit cDNA to mRNA.  
ORGANISM Pyrus communis  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eustosids I; Rosales; Rosaceae; Pyrus.  
REFERENCE  
AUTHORS Abu-Goukh,A.A., Greve,L.C. and Labavitch,J.M.  
TITLE Purification and partial characterization of 'Bartlett' pear polygalacturonase inhibitors  
Physiol. Plant Pathol. 23, 111-122 (1983)  
2 (bases 1 to 1057)  
Slotz,H.U., Powell,A.L.T., Damon,S.E., Greve,L.C., Bennett,A.B. and Labavitch,J.M.  
Molecular characterization of a polygalacturonase inhibitor from Pyrus communis L. cv. Bartlett  
Plant Physiol. 102, 133-138 (1993)  
94151422  
FEATURES  
JOURNAL MEDLINE  
source 1..1057  
/organism="Pyrus communis"  
/strain="Bartlett"  
/db\_xref="taxon:33211"  
/tissue\_type="fruit"  
/dev\_stage="immature fruit"  
34..105  
/gene="PgIP"  
34..1026  
/gene="PgIP"  
/function="precursor of an inhibitor of fungal polygalacturonase"  
/codon\_start=1  
/label=CDS  
/product="polygalacturonase inhibitor"  
/protein\_id="AAB33865.1"  
/db\_xref="GI:169684"  
/translation="MELKFSIFLSLTLFSSVLNPAISDLICNPDDKRVLLQIKKAFG  
PYVLASWKSDTDCDWCVTGDTNTRINSITIFAGOVSGOIPALVGDLPYLETFEH  
KQNLTPGIPQIAIKLKLKFLRSLMTLSSVDFLSQNLKTLFLDSFNLTGAIP  
SSLSOLPMLNALHLDNRKLTHIPKSLQFIGNVDPDLISHNLSGNIPISFAOMDFT  
SIDLSRNKLEGDASVIFGLNKTTQIYDLSRNLEFNLKVEFPTSLTSDINHNKITYG

	gene	SIPHEPOLNQFLMVSYNRLCGIPIVGNKLSQSFDEXSYFHNRCLGABLPSC"	4..
	mat_peptide	/gene="PGIP" 106..1026 /gene="PGIP" /citation=(1) /function="inhibitor of fungal polygalacturonases" /product="polygalacturonase inhibitor"	
	BASE COUNT	268 a 341 c 197 g 251 t	
	ORIGIN		
	Query Match	32.9% Score 328.8; DB 8; Length 1057;	
	Best Local Similarity	60.2%; Pident. No. 6,3e-87;	
	Matches 600; Conservative	0; Mismatches 387; Indels 9; Gaps	3.
Oy	7	atgaatcatcttctgcccattttgtgatatgcagtgttccctcgcctccaac	66
Db	34	ATGMACTCAAGTTCCTCCACCTTCCTCCTTAACCTACTCTTCCTCCGTCAAC	93
Oy	67	ctctcgcatcaagaagtcacaacaacagacagcgttacctaccataaaca	126
Db	94	CCCCTCTTCGCCATCTCTCTCAACCAGCACAAAAGTCTCTCAAAACAANA	153
Oy	127	gccccgaaaaacccccaccattacagactatgggtgtcagaagcagatgttgttg	186
Db	154	GCTTCCGGCGACCCCTACGTTCTTGCCCTATGGAATACTGACACTGCTGCATTGG	213
Oy	187	gacctagtcgaatgtgacgaaccagacaaccgacataattccctataatlaagagac	246
Db	214	TACGCGCTCACTGTGACTCCACCAAAACCGCATTAACCTCCCTACCAATCTTGGCCGGC	273
Oy	247	gaagctctcacgycgaaalcccaacctcagctggtggagactaccaaacctcaagccta	306
Db	274	CAGG--TGTCAAGGCCAATATCCCGCCTTAGTAGAGACTTCCTACTCTGAAACCTT	330
Oy	307	tggctccgaiaaactcccaactcttttcgaaanaalcccaagaanaattctgcactcaaa	366
Db	331	GAATTCATTAAGCAACCCATCTCACTGGCCCAATTCACAAACCGCCATTCGCCAACCTTCAA	390
Oy	367	gacctaaaaaccctcagatcagctcgacagctcagctggccctgccttatcttc	426
Db	391	GGACTCAAGTCTCTGAGGCTCAGCTGAGAACCAACCTTAGGCTCTGCTCGACTTCCTC	450
Oy	427	cctcagcttagaactaacttgtttaactatcgttttaacaacttltyggtyatc	486
Db	451	AGCCAACCTAAACACTTCATCTCCACCTCTCTTCAACACACCTACACGGTGCATC	510
Oy	487	ccctccagcttcccaactcttcggaactttaaagccctgcacttagaagctbaagaaac	546
Db	511	CCGAGCTCGCTTTTGAGCTCCCAAACTCGCGCCTCTTCTCTTAAGCCCAATTAACCTC	570
Oy	547	accggatgaalccccgatatctcttggaatttgcgtgatcc---ccgagacataatct	603
Db	571	ACAGTCAATATTCGATATGCTTTGGCGAGTTCATTGGCAACGTTCCAGACCTGTATCTC	630
Oy	604	tcgcatlaacagctcaaccgggtgtgtcccaaaaacttlctgtagagcagatccaatagg	663
Db	631	TCCCAACAACGCTTTTGATCAATTCACAACTGATTCGATGAGACTTACACAGC	690
Oy	664	ctcgaactcctaaggaaacagactaagaagtgcataatcttctgttggcctaaanaa	723
Db	691	ATAGACTTATACGAGCAACGCTCGAAGGTAGCGCATTCCTGATTTTGGGCGTGAAACAG	750
Oy	724	cgccttgaalaatgcagatttctcaggaagtgcttagtltcaattlcccaaggtgcag	783
Db	751	ACAACCCGAGATTGTGGACCTGTCCAGAACTTCTGGAATTAATGTGTCAAAGT---G	807
Oy	784	gaagttccaccctcttgacataacttagacttgaacataacagatacagcggagctgc	843
Db	808	GAGTTTTCCGACAACCTCGACCTCGTGGATATACACCAATTAAGTCTACGGGAGTATC	867
Oy	844	tcgagtgaaatgagctaaatgcagcactgcagacatttaacgctcagtgataataatcttcg	903

	Db	868	CCAGTGGAGTTTACGCACTGAATTTC	CAGTTTCGTGAACGTAGCTACAAAGCGGTGT	927
Oy	904	gacgaattccaacaggggaaacctccacagatcgcagcgtaagcttatccacaac	963		
Db	928	GGCACATTCCTGTGGGTGGAAGTTGACGAGCCTTGACGAGATATTCTATTTCATTAAC	987		
Oy	964	agtgttgttgytggttcattgccagaatgctg	999		
Db	988	CGATGCTTGTGGGTCTCACTGCCAAGCTGCAG	1023		
RESULT	8				
LOCUS	128276	1058 bp	DNA	PAT	06-FEB-1997
DEFINITION	Sequence 1 from patent US 5569830.				
ACCESSION	128276				
VERSION	128276.1	GI:1819052			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified				
AUTHORS	1 (bases 1 to 1058)				
TITLE	Bennett,A., Labavitch,J.M., Powell,A. and Stoltz,H. Plant inhibitors of fungal polygalacturonases and their use to control fungal disease				
JOURNAL	Patent: US 5569830-A 1 29-OCT-1996;				
FEATURES	Location/Qualifiers				
SOURCE	1..1058				
BASE COUNT	268 a 340 c 198 g 252 t				
ORIGIN	/organism="unknown"				
Query Match	32.9%; Score 328.8; DB 6; Length 1058;				
Best Local Similarity	60.2%; Pred. No. 6,3e-87;				
Matches 600; Conservative	0; Mismatches 387; Indels 9; Gaps				
Oy	7	atgaatacatctctcgtccattttgtgatatagtatgattctctcgtccccaac	66		
Db	34	ATGGAATCAAGTTCTGCACACTTCCTCTCCTAACCTACTCTCTCTCGCTCTTAAC	93		
Oy	67	ctctcgtcaccaaaaagatgcaaacacaacagacaagaagcttaccaccaaaaaa	126		
Db	94	CCGGCTCTCTCGATCTTGCAACCCCGACGAAAAAAGTCTCTTACAANTCAAGAA	153		
Oy	127	gacctgaaaaaccccacatlacagatcatggtgtagtgaagaagatgtgtgttg	186		
Db	154	GCCTTGGGGACCCTTACGTCTTGGCTCATGGAANTCAGACACTGACTGCGCATTTG	213		
Oy	187	gacctgtcgaatgtgagaaacacagcaacgcataattccccataatcaagaagac	246		
Db	214	TACTGCGTCACTGTGATCCACACCAAAACCGCATTAATCTCCTCAGCATCTTGGCGC	273		
Oy	247	gaagcctcaacggccaatccccactcaagtgagggaagacatacctccaagcetta	306		
Db	274	CAG---TCTCAGGCGCAATCCCCTTAGTAGGAGACTTGCCATCTGAAAACCTT	330		
Oy	307	tgtgtcgtlaaatccccaatcttttcggaaaatccacagagaagaatttcgaactaa	366		
Db	331	GAAITTCATAACCAACCAATCTCACTGCCCCATTCACACCCGCCATTGGCCAACCTAAA	390		
Oy	367	gacctaaaatccctcagaactcagctcgacagctcagttgagccctgtcccttaattc	426		
Db	391	GGAGTCAAGTCTTCAGGCTCAGGTGACCAACCTCTCAGGCTGTGTCGACTTCTC	450		
Oy	427	cctcagcttacgaacacacttggttagactatcggtttaacaactttgggtgaatc	486		
Db	451	AGGCAATCAAGAAMCTCACATTCTGTGACCTCTCTTCAACCAACCTCACCGCTGCATC	510		
Oy	487	cctcctcagcttcacactcttcgaaacctlaaagccctcacttaagaagtaagaac	546		
Db	511	CCGAGCTTGTGAGTCCCAAACCTGCGGCTTCTGTTAATCACTCACTCACTCACT	570		

[illegible]

RESULT	9
AF305093	
LOCUS	
DEFINITION	AF305093 1002 bp mRNA PLN 01-MAR-2001 Vitis vinifera polygalacturonase inhibiting protein mRNA, complete cds.
ACCESSION	AF305093
VERSION	AF305093.1 GI:13172311
KEYWORDS	
SOURCE	Vitis vinifera.
ORGANISM	Vitis vinifera
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta Magnoliophyta; eudicotyledons; core eudicots; Vitaceae; Vitis. 1 (bases 1 to 1002) Beziers, A., Lambert, B. and Baillieux, F. Molecular cloning of a grapevine gene coding for polygalacturonase inhibiting protein Unpublished 2 (bases 1 to 1002) Beziers, A., Lambert, B. and Baillieux, F. Direct Submission Submitted (12-SEP-2000) Equipe de Biochimie et Biologie Moléculaire des Plantes (URVVC-EA 2069), Université de Reims Champagne-Ardenne, moulin de la Housse - BP 1039, Reims 51687, France location/Qualifiers
AUTHORS	
TITLE	
JOURNAL	
FEATURES	
Source	1..1002 /organism="Vitis vinifera" /cultivar="Chardonnay" /db_xref="taxon:29760" /clone="7535" /tissue_type="leaves" 1..1002 /codon_start=1 /product="polygalacturonase inhibiting protein" /protein_id="AAK14075.1" /db_xref="GI:13172312" /translation="METSKEFLSSLLVLATPPCSLSERCNPKKKVLTQIKKRR LDPYLIASNNPNTDCGCVEDDITTHRIINSLTIFSGQLSGQIPDAVGLDFLETET IFRKLSTNIGTGPPIPAIAIKLKRLKNRWLSWTFEGVPAPFSBKKULITLDISFNNLGS PIPSLSLTPNLGALHIDRNLTGPIPDSFCKFASTGTGLHSNHLGSKLPYSRSGE
CDS	

BASE COUNT	226 a	324 c	193 g	259 t
ORIGIN	DPRNMDLSRKKLEBGLSIFFNANKSTQIVDSRLFOEDLSRVFEPKSLTSLDLSHKK IAGSLPMMHTSLDLOFLNVSTNRLCGITPVGKQLQSFYDYSIFNNRCLCGAPLQSC"			

Query Match	29.6%;	Score 295.4;	DB 8;	Length 1002;
Best Local Similarity	59.3%;	Pred. No. 5.8e-77;		
Matches 560;	Conservative 0;	Mismatches 376;	Indels 9;	Gaps 3;

[illegible]



	LOCUS	AF159167	909 bp	DNA	PLN	11-APR-2001
	DEFINITION	EucaIyptus grandis polygalacturonase-inhibiting protein (pgip)				
Qy	44	tggtcttccctcgcgccttcccaaacctctcgtcatcatcaagaagtgcacaacaagaagaagc	103			
Db	446	TTATTTTCTTTCTTGCTTGTGCTTCTCTTCACTACTAGTAAATGGCAATCCGAAACACAAA	505			
Qy	104	aagcttctactccaatcaacaacacgccttgcgaataaacccaccatctaacagctcaagggtgc	163			
Db	506	AACTCCTTACAAATAATGAAGAAAGACTTAGGCAATCCTTACCATTTTAGCTTCATGGGATC	565			
Qy	164	cagaagaagcattggttgcgtttggaccctagtcgaatgctgacgaaccaagaacgcgtctaa	223			
Db	566	CAAAACACAAATTGCTGTACTGGTACGTCAATAATGATGACCGGAAACCAACCGGATTA	625			
Qy	224	tttcccttaactcaagaagagaagctctccacggcgcaatccaccctcagaigtgag	283			
Db	626	ATGCTCTCACCGCTCTCCAGACCAA--TATCTCGGCGCAAAATTCGGCGAGCGGTGGAG	682			
Qy	284	acctaccatccatcccaagccttaagtgtcgttcaactccccaatctcttggaaaaatcc	343			
Db	683	ACCTTCATATCTCGAACAACATTGGAAATTTCATCTATTACTATCTCACCGGAAACATTTC	742			
Qy	344	cagaagaatctctgcgactcaagaacctaaatccctcagaactcagctcagccagctctca	403			
Db	743	CACCTGCATATTGCGAAGCTCAACAATCTCAAAAATGTTAAAGCTCAGCTCAGTCACTTA	802			
Qy	404	gtggaccgctgccttattcttcccttagcttgcgaactcgaactgtttagactatcgt	463			
Db	803	CAGGTCCCATCCCTGAAATCTTATAGTACGTGAAGAAATTTGAGCTTGCTCGATGTGAAT	862			
Qy	464	tbaacaaacttttgggttaatccctccctcagcttccactctccgaactcgaaccttaagccc	523			
Db	863	ACAATCAATTTACGGGAACATCCCTCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	922			
Qy	524	tgcacttgaacgttaagaacactcagcggtgaatcccgatctatcttgggaaatttgcgtc	583			
Db	923	TGTACTTGTAGATCGTAACAACATCCACCGGAACATACCGGAATGTTGGAGATTTAAAG	982			
Qy	584	gatcc-----ccgacatatctcttcgataatcaagaactcgaacgggtgttgtcccaaaa	637			
Db	983	GACCAAAATATACCGAGATCTTACTCTTCCACACAACAGTTGACCGGACATGTGCGCGCAT	1042			
Qy	638	cttttgcctagaacagatccaattagctcagctctcgaaggaacagactagaagtgtata	697			
Db	1043	CTTTAGGTGATTTGAATTTTCCACGCTGATTTTCCAGGAATTAAGCTTGAGAGAGATG	1102			
Qy	698	tttcatcttgtttgggcccataaacgccttggaaatgcttagatttttcaggaacgctgc	757			
Db	1103	TTTCGTTTTTGTTCGGGAACAATAAACAAGATCGAGTAATTAATTAATTCGAGGAATTTAT	1162			
Qy	758	ttagtttcaattcttccgaaggtgctcgaaggttttccaccctcttgcataacttagacttga	817			
Db	1163	TGGAAGTTGATATTTCC--GAAATCCGAGTTTCTGTGAGAGCTTGATATCATTTGGATTGGA	1219			
Qy	818	accataaccagatcagcgaagagctctgcagtgaaatttgcgttaacttgcactcagagcat	877			
Db	1220	ATCATATTCGAATTTTGTGGAGCTTACCAACGAGATTAAGAGATTAACCATTTGCAGTTTT	1279			
Qy	878	tttaacgtcagtgtaataatctctgcgcgaagatccaacaggggaaaaacctccagat	937			
Db	1280	TCAATGTGAGATTAAATTAACACTTTGTGGACAAATTCACCAAGGTGAACGTTGCAGAGCT	1339			
Qy	938	tcgaacgtcagcgtctctcccaacaacagttgcttgcgtgcgtcccatctccagaatgct	997			
Db	1340	TTGATATTTTACCTTATTGTCATTAACAATAAGCTTTGTGGCTCTCCCTTCGCGAAATGTA	1399			
Qy	998	ag 999				
Db	1400	AG 1401				

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	AUTHORS	FEATURES	source
AF159167.1	GI:6651275				gene, partial cds.						
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275										
AF159167											
AF159167.1	GI:6651275				</						

```

Db      301 GCTACGCTGGACCACTCTCAGGCTCTGTCTCCTGACTCTCCTCAGCCAACTCAAGAACT 360
Oy      444 aactgtttgactatcggttaacaacttgggtgtatccctccctcaagcttccac 503
Db      361 CACATTCTCTGACCTCTCTTCAACAACCTCACCGGCGCCATCCAGCTCGCTTCTCA 420
Oy      504 tcttcgaacctaaagccctgacttaacgtaacgaactcaccggtgaatcccgga 563
Db      421 GCTCCGAACCTCAAGCTCTTCTATAGACCCCAATMAAGCTCAGAGCTATTCGGA 480
Oy      564 tatcttgggaatttctgtagtc---ccgagacataatcttccgataaccagctcac 620
Db      481 ATCGTTTGGGAGATTTCATTCGCAAGCTTCAAGCTCTATCTCTCCACACAGAGCTTTC 540
Oy      621 cgggttcttcccaaaccttctgtagagacgataccaattagctcgaactctcaaggaa 680
Db      541 GGGCAACATTCACACCTCTATTGCCCAGATGACCTTGGCGAAGCATATACACGAA 600
Oy      681 cagaatagaagtgatattcatcttctgttggcccttaaaaaagcttggaaatgtctaga 740
Db      601 CAAAGCTCGAGACGCA---TCAGTGATATTGGGCTGAACAAGAACCCCAAGATTGTGA 657
Oy      741 ttttccaggaacgtagttagttcaattctccaggggtgcaggaagttccaccctctt 800
Db      658 CCTATCCAGAGACTTGCTGGAATTATCTGTCAAGGT---GAGATTTCGACAAGCTT 714
Oy      801 gacatactagacttgaacataaccagatcagcgggaagctctgtcgaagtgaattgctaa 860
Db      715 GACCTCAGCTGGATGTAACACACATAGATCTACGGAGATATCCAGTGGATTACCA 774
Oy      861 attgagacctgacatlaaacgtatgaataatctctcgcgcgaagaatcccaacagg 920
Db      775 ACTGATTTCCAGTTCCTGTAACGTACGTAACAGCGCTGTGTGATTCAGATTCAGTGG 834
Oy      921 gggaacctccagagatcgaacgcgtacgacctatctccacaacagttgctgtggtgc 980
Db      835 CGGAAGTTGCAAACTTCAACGAGTATTCTTATTTCCATAACCATGCTGTGTGTGC 894
Oy      981 tccattgc 988
Db      895 ACCCTCC 902

RESULT 13
AF159170 909 bp DNA PLN 11-APR-2001
LOCUS Eucalyptus saligna polygalacturonase-inhibiting protein (pgip)
DEFINITION gene, partial cds.
ACCESSION AF159170
VERSION AF159170.1 GI:6651281
KEYWORDS Eucalyptus saligna.
SOURCE Eucalyptus saligna.
ORGANISM Eucalyptus saligna.
REFERENCE 1. Eukaryota: Viridiplantae. Streptophyta. Embryophyta. Tracheophyta.
AUTHORS Rosidae: eucosids II: Myrtales; Myrtaceae; Eucalyptus.
TITLE Chlamydomonas, P.M., Botha, A.M., Wingfield, M.J. and Wingfield, B.D.
JOURNAL Molecular relatedness of the polygalacturonase-inhibiting protein
REFERENCE Theor. Appl. Genet. 102, 645-650 (2001)
AUTHORS 2. (bases 1 to 909)
JOURNAL Chlamydomonas, P.M., Oberholster, A., Wingfield, M.J. and
REFERENCE Wingfield, B.D.
TITLE Direct Submission.
JOURNAL Submitted (15-JUN-1999) Genetics, Faculty of Agricultural and
AUTHORS Biological Sciences, University of Pretoria, Forestry and
REFERENCE Agricultural Biotechnology Institute, Pretoria 0001, South Africa
FEATURES
Source
1. 909
/organism="Eucalyptus saligna"
/db_xref="taxon:99019"
<1..>909
mRNA

```

```

gene /gene="pgip"
      /product="polygalacturonase-inhibiting protein"
      <1..>909
CDS /gene="pgip"
     /codon_start=2
     /evidence=experimental
     /product="polygalacturonase-inhibiting protein"
     /db_xref="GI:6651281"
     /translation="EFNPDDKVLVLIQIKAFEDPVLVSKMSDIDCCWCVTCOSTT
     NRINSLTIFACQVSGQIPALVGDLPYLETFEHHKQNPJTGPIQPAIAKLGIFLRIS
     WTNLSGVPDEFSLQKLNLFDLSENNLTGAIPLSSLSQPLNHLDRNKLTGHI PK
     SEGQPIGVNPDLYLSHNLGNIPTPSFQMDGKRLSRNKLDEASVIFGLNKTQIV
     DLSRLLEFNLSKVEFPTSLTSLDYNHKKIIGSIFVETQNLNFORLANSYRNLQIQIP
     VGRKIOSNEYSYFNRLCLCAPLHTA"
BASE COUNT 236 a 286 c 178 g 209 t
ORIGIN

Query Match 28.4% Score 284; DB 8; Length 909;
Best Local Similarity 60.1%; Pred. No. 1,4e-73;
Matches 546; Conservative 0; Mismatches 350; Indels 12; Gaps 4;

Oy      84 atgcaacaacaacgaacgaacgaacttactccaatacaaacgaacttgaataaccac 143
Db      4 ATTCAATCCGAGACAGAAAGAGTCTCTTCAATCAAGAAAGCTTGGCGAGCCCTTA 63
Oy      144 catlcagaactcattggtgtcagacgaagatgtgtgtgtggaactagtcgaattga 203
Db      64 TGTCTTGGCTCTCATGGAATACAGACACGAGCTGTGTGATTGTGACTGCTGACCTGTGA 123
Oy      204 cgaacacgaacccgataattccctctataatcaagaacgaacgtctcaccggcca 263
Db      124 CTCACCCCAAAACCGCATCAACTCCCTCACCATCTTTCGCCGCCAGG---TATCCGGCA 180
Oy      264 aatccacactcagtggtgagacactacacataccccaagccttagtgcgtaacacc 323
Db      181 AATCCCGCCCTAGTTGAGAGACTTCCCTTGAACCTTGATTCACACAGCAAC 240
Oy      324 caatcttgggaanaatcccaagaagaattctcgaactcaagaacaaatccctcag 383
Db      241 CAATCTCACTGAGCCCAATCCCAACCGCATTCGCCAAGCTCAAGAGCTCAAGTTCTCAG 300
Oy      384 actgaactcgaacgtctcgaatggtccctgtcccttatcttccctcaagctcagaact 443
Db      301 GCTCAGCTGGACCAACCTCTCAGGCTGTGCTGACTTCTCAGCCCAACTCAAGAACCT 360
Oy      444 aactgtttagactatcggttaacaacttgggtgtatccctccctccctcaagcttccac 503
Db      361 CACATTCTCTGACCTCTCTTCAACAACCTACCGGCGCCATCCCAAGCTCGCTTCTCA 420
Oy      504 tcttcgaacctaaagccctgacttaacgtaacgaactcaccggtgaatcccgga 563
Db      421 GCTCCGAACCTCAAGCTCTTCTATAGACCCCAATMAAGCTCAGAGCTATTCGGA 480
Oy      564 tatcttgggaatttctgtagtc---ccgagacataatcttccgataaccagctcac 620
Db      481 ATCGTTTGGGAGATTTCATTCGCAAGCTTCAAGCTCTATCTCTCCACACAGAGCTTTC 540
Oy      621 cgggttcttcccaaaccttctgtagagacgataccaattagctcgaactctcaaggaa 680
Db      541 GGGCAACATTCCAACCTTATTCGCAAGCTTCAAGCTCTATCTCTCCACACAGAGCTTTC 600
Oy      681 cagaatagaagtgatattcatcttctgttggcccttaaaaaagcttggaaatgtctaga 740
Db      601 CAAAGCTCGAGACGCA---TCAGTGATATTGGGCTGAACAAGAACCCCAAGATTGTGA 657
Oy      741 ttttccaggaacgtagttagttcaattctcgaagggtcgaagatctcaccctctt 800
Db      658 CCTATCCAGAGACTTGCTGGAATTATCTGTAAGGT---GAGATTTCGACAAGCTT 714

```



	RESULT	14	
	AF159171		
	LOCUS	AF159171	
	DEFINITION	Eucalyptus nitens polygalacturonase-inhibiting protein (pgip) gene,	
	ACCESSION	partial cds.	
	VERSION	AF159171	
	KEYWORDS	AF159171.1 GI:6651283	
	SOURCE	Eucalyptus nitens.	
	ORGANISM	Eucalyptus nitens. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Myrtales; Myrtaceae; Eucalyptus. 1 (bases 1 to 909) Chlwmamurcombe,P.M., Botha,A.M., Wingfield,M.J. and Wingfield,B.D. Molecular relatedness of the polygalacturonase-inhibiting protein genes in Eucalyptus species Theor. Appl. Genet. 102, 645-650 (2001)	
JOURNAL	TITLE	Chlwmamurcombe,P.M., Oberholster,A., Wingfield,M.J. and Wingfield,B.D. Direct Submission Submitted (15-JUN-1999) Genetics, Faculty of Agricultural and Biological Sciences, University of Pretoria, Forestry and Agricultural Biotechnology Institute, Pretoria 0001, South Africa	
REFERENCE	AUTHORS	Location/Qualifiers 1..909 /organism="Eucalyptus nitens" /db_xref="taxon:87674" <1..>909 /gene="pgip" /product="polygalacturonase-inhibiting protein" <1..>909 /gene="pgip" <1..>909 /gene="pgip" /codon_start=2 /evidence="experimental /product="polygalacturonase-inhibiting protein" /protein_id="AAF22252.1" /db_xref="GI:6651284"	
FEATURES	source	/translation="EFNPDKQVLLQIKKAFGDPEYLASWSKSDTDCDWCVTCDSST NRINSLTIFAGVSGEIPALVGDLPLYELETFHKOPIITGPIDPAIAKLKLFRLS WTNLSGVDPIFSOLKMLTEFLDLSFNNLTGAIPSSLSQPLNLHLDRNKLTGHIP SFGGEIIVNPDLVLYSHNOISGNIPTSPFAOMDFGKHRSRNKLGDAVSFGUKNTQIOI DSGNLIEFPNLKVEPFTSLISLDVNHNKIIGTSIPVERTQLNQPLANYSNRLCQIRI VGGLQSGFNESYFRHNRCGLCAPLRATYA"	
BASE COUNT	CDS	236 a 285 c 179 g 209 t	
ORIGIN			

Query Match	28.4%	Score 284;	DB 8;	Length 909;
Best Local Similarity	60.1%	Pred. No. 1,4e-73;		
Matches 546; Conservative	0;	Mismatches 350;	Indels 12;	Gaps

Db	ATTCAATTCGGAGCAAGAAGGCTCTCTACAAATTAAGAAAGAGCTTCGGCCAGCCCTTA	63
Qy	144 catcaagactcattcaggtctgcaagcagcgttgttgggttggagcttagtcgaatgta	203
Db	64 TATCTTGGCTCATGGAATTCAGACACCGACTGTGTGATTGGTACTGGCTCACTGTGA	123
Qy	204 cgaacacagcaacgcgcaataatccctcataatltcaagaagcaagcgtctccagcgca	263
Db	124 CTCAACCAACAAACCGCATCAACTCCCTCACCACTCTTTGGCCGCCAGG--TATCCGGCGA	180
Qy	264 aatcccaactcagttggaggagaccataacctccaaagccttatgtgttcgtaaatccc	323
Db	181 AATCCCGCCCTTAATTGGAGCTTGGCCGTGACTTTGAAACCTTTGAATTCACAGCAACC	240
Qy	324 caactcttcggaanaatcccaagaanaatlttcgcataaagacctaataatccctcag	383
Db	241 CAATCTACTGTGCCCATTCACACCGGCATTTGGCAACTCAAGAGACTCAAGTTCTCAG	300
Qy	384 aactcagtcgacagtcctcagttgcccgtctcccttatcttccctcagcttgcgaact	443
Db	301 GCTCAGCTGGGACCAACTCTCAGGCTGTGCCGACTTCTCAGCCAACTCAAGAACGT	360
Qy	444 aactgtttgagactatcglttcaaaactttggggtgtaactccctccagctttccac	503
Db	361 CACATTCCTGACCTCTCTCTTCAACAACGTACACGGGCGATCCCAAGCTGCTTTCTCA	420
Qy	504 tcttcgaaaccttaagccctcgaccttagaagcgtlaaagaaactcagcgtgaaatcccgga	563
Db	421 GCTGCCAAACCTCAACGCTGTCTATCTAGAACCGCAATTAAGCTACAGGTGATATTCGAA	480
Qy	564 tatctttgggaatttgcctgatcc---ccggaatatatcttgcgataaacagctcac	620
Db	481 ATCTGTTTGGGAGATTCAATTGGCAAGTTCCAGACTGTATCTCTCCACAAACGCTCTC	540
Qy	621 cgggtttgttcctcaaaactttgcttagagcagatccaaatlagcctcgacttccaggga	680
Db	541 GGGCAACTTTCACACTCATTGTGGCCAGATGAGACTTGGCAAGCATTAAGATATCAGGAA	600
Qy	681 cagaactgaagtgatattcatctctgtttggccttaaaaaagcgtttgaaatgctaga	740
Db	601 CAACCTC---GGGAGCACCATCAGATTAATTGTGGCTTAACAGCAACCCAGATTGTGGA	657
Qy	741 ttttcaggaagtgcttagttcattcattcccaagggtcagagttccaccctctt	800
Db	658 CCTATCCAGAACTTGTGCAATTAAATCTGTCTCAAGCT---GGAGTTTCCGACAAGCTT	714
Qy	801 gacatacttgaacttgcacataacacagacacgacgagagtcgtgcagtgaaatltgctaa	860
Db	715 GACTCACTGATGATTAACCCACATTAAGATCTAAGGAGATATCCCAATGAGATTACCA	774
Qy	861 attgagcctgcagacatttaacgttcagtgataataaactctctggcgcaagatltccaacag	920
Db	775 ACTAATATTTCCAGTTCTCTGAACGGACTCAACAGGCTGTGGTAGATTTCCAGTGGG	834
Qy	921 gggaaaccttcagagatcgcagcgtlaecgctatctccacaagaagtgtcgtgtgtgc	980
Db	835 CGGAAAGTTGCAAAAGCTTCAACGAGTATTCTTATTTCATTAACCGAATGCTGTGTGTGC	894
Qy	981 tccattgc 988	
Db	895 ACCCCTCC 902	

CSPGIP	CSPGIP	1207 bp	mRNA	PLN	09-OCT-1996
LOCUS					
DEFINITION	C.silensis mRNA for polygalacturonase-inhibiting protein.				
ACCESSION	Y08618				
VERSION	Y08618.1	GI:1617033			
KEYWORDS	polygalacturonase-inhibiting protein.				
SOURCE	Citrus sinensis.				
ORGANISM	Citrus sinensis				

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;  
Rosidae: eucosids II; Sapindales; Rutaceae; Citrus.

REFERENCE 1 (bases 1 to 1207)  
AUTHORS Mayer, R.T.  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1207)  
AUTHORS Mayer, R.T.  
TITLE Direct Submission  
JOURNAL Submitted (03-OCT-1996) R.T. Mayer, USDA, ARS, HRS, Subtropical  
Plant Insect Research Unit, 2120 Camden Road, Orlando, Florida,  
32803, USA

FEATURES  
source location/Qualifiers

1..1207  
/organism="Citrus sinensis"  
/cultivar="Hamlin"  
/db\_xref="taxon:2711"  
/dev\_stage="embryonic"  
/tissue\_type="callus"  
14..997  
/codon\_start=1  
/product="polygalacturonase-inhibiting protein"  
/protein\_id="CA69910.1"  
/db\_xref="GI:1617034"  
/db\_xref="SPTREMBL:O96401"  
/translation="MSNTSLSPFFLICISPSLSDLCNPNDKXVLLPKKSLNPNY  
VLASNPRTDCCWYCTCDLTNNRINSLTIRAGDLPGLPPVGGGLPIETLMPKL  
PSLTGPIQPIAIKLNKLTIRISWINSIPSPVDFIQLTLEFNNLSQTLIGS  
LSRLQGLALHIDRNLTGSIPESEFTGSIIPDLXLSHNOJSGKIPASIGSDFTMI  
DISRNKLEGDASFELNKTORTIDVSRNLEERLSEKVEFSPITNLDNHNKIFGSI  
PAQITSLNENGLFVNSYRNLCPDPVGGKLOSFGYEFYFHNRLCGAPLER"

BASE COUNT 331 a 279 c 239 g 358 t  
ORIGIN

Query Match 28.4%; Score 284; DB 8; Length 1207;  
Best Local Similarity 59.4%; Pred. No. 1.5e-73;  
Matches 558; Conservative 0; Mismatches 370; Indels 12; Gaps 4;

QY 66 cctctctgcatcaacaagatgcaacaagaagaagcttactccaatcaaac 125  
DB 67 CCCTTCACTCTCAAGACTCTCAACCAATGACAAAGAGTCTCTCAAAATTCAAAA 126  
QY 126 agccttgaataaaccccaacttaagaactcattggtgagcagaagattgtgtgtg 185  
DB 127 ATCTTTGAACACCTTACCTTACCTTCTTGAACCCAAAAAAGTGTGACTG 186  
QY 186 ggaactagtcgaatgtgaagaacacgacacataatccctcctaataatcaagcga 245  
DB 187 GTACTGGGTACATGCATCTCACACTAACCGCATCACTCTCTACCAT---CTTCGC 243  
QY 246 cgaagcttcaacgagcaaatcccaactcagtggtgagacctacacatcccaagcct 305  
DB 244 CGGAGATCTCCCGGCGAGATCCCGCCGAAAGTTGGATCTTCTTCACTCGAAAAACCT 303  
QY 306 atggttcgaaatccccaactcttccgaaaaatcccaagaagaatttcgacacaa 365  
DB 304 AATGTTTACAAAGTACCCAGCTCACTGGCCCAATACAAACCGCATTTGCCAAGCTCAA 363  
QY 366 agacctaaatccctcagactcagctcgacagatctcagtgagccctgtcccttatctc 425  
DB 364 AAACCTGAACAGCGCTAGCTTAGCTGGACAACATTTCTGGGCGGTTCTGATTTTAT 423  
QY 426 ccctcagactcagaataactgttagactatcggttaacaactttgggtgtaact 485  
DB 424 CAGGCAACTCACCAACTTACATTTCTGGAGCTTTCATTTAACAAATCTTCTGGAGAGAT 483  
QY 486 cccctccagcttcccaacttccgaaccttaagccctgacacttaagaagcgaagc 545  
DB 484 CCCAGGTTCACTTTGAGAGTGCAGAGCTTGGCGCTTTTCATTTTGACAGAAACAGCT 543  
QY 546 caccggtgaatcccccatalctttgggaatttgcgt---gatcccgagacatalact 602

DB 544 TACGGGTTCAATCCCGAGTCTTTTGSTACATTCACCGGAGATATACCTGATCTTACTT 603  
QY 603 ttgcataaaccaactcaccggttgttcccaaaactttgtcagagacagatccaaatag 662  
DB 604 GTCAATTAACCACTCTCTGCAAAATCCCTGCTCTTTAGGACAGATGGATTTTAACAC 663  
QY 663 gctcgaacttccaggaacagactagaaggtgatattcattcattctgttggccttaaaa 722  
DB 664 CATTGACTTCTCAGGACAAAGCTTTGAGGCGATGCTTCTTGTGTCGGGTTGAACAA 723  
QY 723 acgcttgaataatgtagatttccaggaacgagtgctagttcaatttcccaaggttgc 782  
DB 724 GAGAGCGACAGAAATGATGTTTCAAGAAACTGTGGATTTAATCTGTCCAAGT--- 780  
QY 783 ggaagttccacccctttagacataacttagacttgaaccataacacagatcagcgaagtc 842  
DB 781 TGAGTTTCCACAGAGCTTGAACAAATTTGGATTGAAATCACACAGATATTTCCGGAGCAT 840  
QY 843 gtgcagtgaaatggct---aattggaacctgcagacatttaacgtcagtgataataatct 899  
DB 841 TCCTGCTCAATTAATTCTTCTGCTACAGAAATCTAGCATTTCTTAATGTCAGTTACACAGTT 900  
QY 900 ctgcggaagaatcccaacagggggaacacccacagagatcgcacggtacggcctatccca 959  
DB 901 GTGTGGGCGCATTCCTCGTGGGGGAAAGTTGCAGAGCTTTGATATACGAGATATTTTCA 960  
QY 960 caacagttgcttgtgtgtgtcctcatatgccaagaatgctag 999  
DB 961 TAATAGTGCTTGTGTGCGCCGCCCTCGAAGATGAAG 1000

Search completed: March 13, 2002, 20:12:55  
Job time: 4596 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2002, 19:33:49 ; Search time 103.51 Seconds  
(without alignments)  
8274.253 Million cell updates/sec

Title: US-09-308-140-6

Sequence: 1 atgataatcattcaatcattc.....ctccattgcagaatgctag 999

Scoring table: IDENTITY-NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

N.Geneseq.1101.\*  
1: /SID52/gcgdata/geneseq/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	328.8	32.9	1058	17	AAT49434
2	290.4	29.1	2075	17	AAT49435
3	256.8	25.7	1122	21	AAC48852
4	253.4	25.4	1132	21	AAC33037
5	250.8	25.1	1164	21	AAC33759
6	249.2	24.9	1157	21	AAC48837
7	218	21.8	1175	21	AAC39926
8	194.8	19.5	1079	21	AAC36540
9	181	18.1	1518	21	AAC54534
10	177.6	17.8	1528	21	AAC52792
11	172.4	17.3	1117	14	AAO42595

12	170.8	17.1	2917	18	AAT86755	Raspberry dru1 gen
13	138.8	13.9	792	14	AAO42596	Sequence of varian
14	97.6	9.8	1400	22	AAO50107	Maize disease resi
15	75.6	7.6	3341	19	AAV14522	CF-5 pathogen resi
16	71.8	7.2	3979	19	AAV14518	CF-5 pathogen resi
17	71.8	7.2	3979	19	AAV14519	CF-5 pathogen resi
18	71.8	7.2	4123	19	AAV14523	CF-5 pathogen resi
19	68.8	6.9	3573	17	AAT06307	Partial tomato pat
20	68.8	6.9	6471	17	AAT06306	Tomato pathogen re
21	62.6	6.3	3396	17	AAV79268	Pinus radiata cell
22	60.2	6.0	3176	18	AAT62124	Pinus radiata cell
23	59.6	6.0	1052	21	AAV79691	Pinus radiata cell
24	59.4	5.9	936	22	AAV58252	Pinus radiata cell
25	59.4	5.9	936	22	AAV58254	Oligonucleotide D1
26	59.4	5.9	936	22	AAV58257	Oligonucleotide D1
27	59.4	5.9	936	22	AAV58259	Oligonucleotide D1
28	59.4	5.9	936	22	AAV58262	Oligonucleotide D2
29	59.4	5.9	936	22	AAV58255	Oligonucleotide D2
30	59.4	5.9	943	21	AAC39834	Oligonucleotide D1
31	59.4	5.9	1208	21	AAC40606	Arabidopsis thalia
32	59	5.9	784	21	AAC41583	Arabidopsis thalia
33	59	5.9	810	21	AAC45806	Arabidopsis thalia
34	58.6	5.9	3552	21	AAC44718	Arabidopsis thalia
35	57.8	5.8	1204	21	AAC46299	Arabidopsis thalia
36	57.8	5.8	2175	21	AAV79686	Arabidopsis thalia
37	56.8	5.7	936	22	AAV58252	Oligonucleotide D1
38	56.8	5.7	936	22	AAV58254	Oligonucleotide D1
39	56.8	5.7	936	22	AAV58257	Oligonucleotide D1
40	56.8	5.7	936	22	AAV58259	Oligonucleotide D1
41	56.8	5.7	936	22	AAV58262	Oligonucleotide D2
42	56.8	5.7	1259	21	AAV58255	Oligonucleotide D2
43	56.6	5.6	4017	21	AAC39335	Arabidopsis thalia
44	56.2	5.6	4036	21	AAV79728	Arabidopsis thalia
45	56.2	5.6	4036	21	AAV79681	Pinus radiata cell

#### ALIGNMENTS

##### RESULT 1

ID AAT49434 standard; CDNA; 1058 BP.

AC AAT49434;

DT 05-MAR-1997 (first entry)

DE Pear polygalacturonase inhibitor protein cDNA.

XX PCIP: polygalacturonase inhibitor protein; fungal resistance;

KW Botrytis cinerea; transgenic plant; tomato; strawberry; ss.

XX Pyrus communis cv. Bartlett.

OS Key Location/Qualifiers

FT CDS 34..1026 /tag= a

FT /product= polygalacturonase\_inhibitor

PN US5569830-A.

XX 29-OCT-1996.

PF 03-MAY-1994; 94US-0238163.

PR 03-MAY-1994; 94US-0238163.

PA (REGC ) UNIV CALIFORNIA.

PI Bennett A, Labavitch JM, Powell A, Stoltz H;

DR WPI, 1996-496968/49.

DR P-PSDB: AAW09095.

*Hand done*

XX DNA construct, comprising polygalacturonase inhibitor protein DNA  
 PT - useful to confer resistance to fungal, partic. Botrytis cinerea,  
 PT infection in plant, pref. tomato or strawberry  
 XX  
 PS Claim 2; Fig 1; 33pp: English.

CC The pear polygalacturonase inhibitor protein cDNA was obtained using  
 CC mRNA isolated from mature green fruit. Pear PGIP exhibits differential  
 CC inhibition of polygalacturonases from different fungal species:  
 CC introduction of the pear PGIP cDNA into other plant species alters the  
 CC susceptibility of these plants to the fungal pathogens against which  
 CC pear PGIP is effective. For example, transgenic tomato fruit  
 CC expressing high levels of pear PGIP were shown to be more resistant  
 CC to the fungus Botrytis cinerea than control fruit.  
 XX  
 SQ Sequence 1058 BP: 268 A; 340 C; 198 G; 252 T; 0 other;

Query Match 32.9%; Score 328.8; DB 17; Length 1058;  
 Best Local Similarity 60.2%; Pred. No. 1.7e-96;  
 Matches 600; Conservative 0; Mismatches 387; Indels 9; Gaps 3;

OY 7 attgatacatcttctcgcctatttctgcatatgcatgatttctctcgtcccaaac 66  
 DB 34 atggaactcaagttctcaccctctcctcctcaaccctactctctcctcgctctaac 93  
 OY 67 ctctctgcatcacaagaatgcaacaagaagaagaagttacttcccaataaaca 126  
 DB 94 ccgctctctcgtatctctcgtacacccgacgaacaaagccctccataaataa 153  
 OY 127 gctctgaaaacccacacatatacagactatgggtgcagacagattgtgtgttg 186  
 DB 154 gctctcgagcaccctcactgcttctgctcattggaatcagacatgctgtgatg 213  
 OY 187 gacctagtgaaatgtgcagaaaccgaacagcatttccctcataatcaagacgac 246  
 DB 214 taacgctcaccgttgatctccacacacacacgaacttaccctccacacatcttgcg 273  
 OY 247 gaagctcaccgccaacatccacacacacacacacacacacacacacacacac 306  
 DB 274 cagg---tgacagggcacaatcccgccctagttagagacttgcacacacacacac 330  
 OY 307 tgggtccgttaactcccaactcttctcgaaaataccacagaagaattctgcac 366  
 DB 331 gaattccataagaacccaacttccactgacccaatccacacacacacacacacac 390  
 OY 367 gaactaaatccctcagacacacacacacacacacacacacacacacacacacac 426  
 DB 391 ggaactcaagctctcagctcagctcagctcagctcagctcagctcagctcagct 450  
 OY 427 cctcagctcagaacttaacttctgtagactatcgcttgaacttctggtgaac 486  
 DB 451 agccaactcaagaactcactcactcactcactcactcactcactcactcactcact 510  
 OY 487 cctcctcagctcactcactcactcactcactcactcactcactcactcactcact 546  
 DB 511 cccaagctccttctcagctcactcactcactcactcactcactcactcactcact 570  
 OY 547 accgggaatcccgacatcttctggaattctgtcgatc---ccgagacataact 603  
 DB 571 acaggtcactatccgacatcttctggaattctgtcgatc---ccgagacataact 630  
 OY 604 tcgataaccacagctcagcggttcttcccaaaacttctcagagacacacacacacac 663  
 DB 631 tcccaacacagcttctgtaataatccacacacacacacacacacacacacacacac 690  
 OY 664 ctcgacttctcagagac 723  
 DB 691 atagacttatacagcgaagaagctgaaggtgacacacacacacacacacacacac 750  
 OY 724 cgcttgaaatgtagatttctcaggaacgctgtagtttcaattctccaggtgacg 783

DB 751 aacaccagatttgagacctgcaggaacttgcgtgaatttaactcgtcaaaagt---g 807  
 OY 784 gatttccacctcttctactacttaacttgaacccaacacagatcagcgagctcg 843  
 DB 808 gatttccgacaagcttgcacctggtgatatcaaccacacataagatctacggagatc 867  
 OY 844 tcgagtgatttggctaaatttgagacctgcagacatttaacgctcagtgataatctctgc 903  
 DB 868 ccaatggagatttaacgaacttaattccagttccttgaacgctgaacacacagcgctgt 927  
 OY 904 gccaagatttccacaagggggaacacctccagagatcagcgtacgcctatccacacac 963  
 DB 928 gtcagatttctcgtggtggaagttgcagagccttcagcagatcttcttattccataac 987  
 OY 964 agtgccttggtgtgtctccatctgcaggaatgctag 999  
 DB 988 cgatccttggtggtgtctccatccactcccaagctgcaag 1023

## RESULT 2

AAT49435 standard; DNA: 2075 BP.

XX AAT49435;  
 AC  
 XX  
 DT 05-MAR-1997 (first entry)

DE Tomato polygalacturonase inhibitor protein genomic DNA.

XX PGIP; polygalacturonase inhibitor protein; fungal resistance;

KW Botrytis cinerea; transgenic plant; tomato; strawberry; ss.

XX Lycopodium esculentum.

XX Key Location/Qualifiers

FT CDS 421..1404  
 FT /tag= a  
 FT /product= polygalacturonase\_inhibitor  
 FT 1545..1550  
 FT /tag= b

PN US5569830-A.

XX 29-OCT-1996.

XX 03-MAY-1994; 94US-0238163.

XX 03-MAY-1994; 94US-0238163.

XX (REGC ) UNIV CALIFORNIA.

XX Bennett A, Labavitch JM, Powell A, Stoltz H;

XX WPI: 1996-496968/49.

XX P-PSDB: AAM09096.

PT DNA construct, comprising polygalacturonase inhibitor protein DNA  
 PT useful to confer resistance to fungal, partic. Botrytis cinerea,  
 PT infection in plant, pref. tomato or strawberry

PS Claim 4; Fig 2; 33pp: English.

CC A fragment of tomato genomic DNA without introns and coding for  
 CC the polygalacturonase inhibitor protein was isolated from mature  
 CC green fruit. The DNA coding for PGIP can be introduced into other  
 CC plant species, e.g. strawberry, to alter the susceptibility of  
 CC these plants to the fungal pathogens against which tomato PGIP is  
 CC effective.

SQ Sequence 2075 BP: 660 A; 366 C; 348 G; 701 T; 0 other;

Query Match 29.1%; Score 290.4; DB 17; Length 2075;

Best Local Similarity 58.6%; Pred. No. 7.5e-84;  
Matches 564; Conservative 0; Mismatches 386; Indels 12; Gaps 3;

```
OY 44 tgaatttccttgcttccaaacctctgcatcacaagaatgcaacaacagcagaac 103
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 446 ttaatttcttgcttgcttccctcaactacagtaagatgcaatcghaagacaaa 505
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 104 aagcttactccaacaacagccttgaanaaccccccataagactatggtgt 163
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 506 aagctcttcacaanaaagaagacttaggcattccattagcttcctggatc 565
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 164 cagaacagcattgtgtgttggaacctagtcgaatgtgcaaaccaacgcatata 223
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 566 caaacacagattgctgttactgtgtaatgataaatgtgacggaataacacagcatata 625
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 224 ttccctcaaatccaagcagcaagctctcacgagccaatccacactcaggtggag 283
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 626 atgcttcacgctctccacagcaca---tatctcgccaattccgagcgcgtcgag 682
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 284 acctaccatacctccaagccttatgtttccgtaaacctcccccacttttcggaataatcc 343
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 683 acctccataatctgaaacattggaattccaatgtaactaattccacacggaacaattc 742
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 344 cagaagaanaattctgcatcacaagacttaaaatccctcagactcagctcagactctca 403
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 743 caacttgaatttggaagctcacacaatccaataatgttaaggtcagcttcacataactta 802
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 404 gtggccctgtcccttataatcttcctcctcagcttaacgaactatgtttagacttcgt 463
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 803 caggtccgactccctgacttctcctagtcagctgaagaattgacgtgtgcgttgaatt 862
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 464 ttaacaaacttttggttgaatccctccctcagcttccactctccactccgaacctaagccc 523
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 863 acaataaattcacggacaacatccctctccctccctcctaagcttccgaaattgcatcga 922
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 524 tgcacttaagaacgtaacgacactcaccggtggaatcccgatattcttgggaatttgcgcg 583
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 923 tgaatttatgctgaacaaactcacacggaacataaccggaatcgtttggagatttaag 982
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 584 gatcc-----ccggaatataatcttgcataacacagctcaccgggttctgtcccaaaa 637
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 983 gaacaaataatacagatctctacacttccacacacagcttgacccgagacatgtgcgcgat 1042
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 638 ctgttgctagcagatccaattaggtcgtgacttccaggaacagacagaagtgata 697
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1043 ccttagtgtaatttgaaatttccacgctgacttccccaaggaataagcttgaaggagatg 1102
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 698 ttcatctctgtttggccctaaanaacgcttgaatgtctagatatttccaggaacgctgc 757
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1103 ttctgttttctgtcggaagaataagacagctcaggttaattgattatcagaagatttat 1162
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 758 ttagttccaatttccacaggtgcaggaatttccacccctcttgcacatcttagacttga 817
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1163 tggagtttgatatctc---gaaatcggaatttgcgaagcttgatataatcattgatttga 1219
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 818 accataaacagatcagcggaagtctgtcagtgaaatgtgctaatttgcacccgcgcgaact 877
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1220 atcataaattcgaatttctgttagcttaccacccaggaattgaaagatgtaccactcgatttt 1279
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 878 ttaacgtcagctataataatctctcgcgcaagaatcccaacaggggaaccccgagat 937
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1280 tcaatgtgagttataatagacttttgcacagatcccaaaaggttgaacggttcgcagact 1339
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 938 tgaacgttaagcctatactccacaacagttgtgtgtgtgtcctcatgtcagaatgct 997
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1340 ttcataatttactctatttgcataacaatgaccttgcgtccctccctgcagaaatgta 1399
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 998 ag 999
    | |
DB 1400 ag 1401
```

```
AAC48852
ID AAC48852 standard; DNA; 1122 BP.
XX
AC AAC48852;
XX
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 59011.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
XX
PD
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 17-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
```

```
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144682.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 27-AUG-1999; 99US-0151068.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 25-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 25.7%; Score 256.8; DB 21; Length 1122;
Best Local Similarity 56.0%; Pred. No. 4.6e-73;
Matches 549; Conservative 0; Mismatches 422; Indels 9; Gaps 3;

QY 26 ctatttgcataatgatatttctctcgtccccaacccctcgcatacaagaat 85
Db 37 cgacatgtgtctcttctgttcttgcattctccacgacctgttgtcctaagatctc 96
QY 86 gcaacaacaagcaagcaagccttactccaatccaacaagccttgaaaacccacca 145
Db 97 gtaacccaatgtacaacaaacacccctcctaagatcaagaatcttccaacacccctatc 156
QY 146 ttaacagactatgggtgtcagacagacgatgtgtgttgggaacctagtcgaatgtgacg 205
Db 157 acctgcctcatgtgacccccaacagacgttgcctcctgtactcgttgggtgcgagc 216
QY 206 aaaccagc---aacgcgataatttccctcataataatcaagaagcagacgtccacccgagc 262
```

Db 217 acgcccacgttaaccacgctgttaacgcgttaacatatcttcgcgcagatctccggtc 276  
Oy 263 aatccaccctcaggttgaggagactaccataccccaagccttatggttcgtaactcc 322  
Db 277 agatcccggtcgtgaagtcggttgcgtccgtatcttgagaccccttgcctcgaactc 336  
Oy 323 ccaatccttctggaaaataccagaagaattctgcactcaagaacctaaatccctca 382  
Db 337 ctaactcaaccggttaacatccaccacatcgcaactcaagaacctcgcgaatgctca 396  
Oy 383 gactcagctcagaccagctcagtgccgttcgtcccttatcttcctcctaagctaaac 442  
Db 397 ggtcagctcgtgaacgagatcgtgaaggtccaatcccgacttataagtcagccaagatc 456  
Oy 443 taacttgctgacttaactcggttaacaaacttgggtgtaatccctccctcagctcca 502  
Db 457 tgcagttcttgaacttccctcctcaatgtctctcgttgcattcacaagtctctcctca 516  
Oy 503 ctcttcggaaccttaagccctgcacttaagaacgtaacgaactccggtgaatcccg 562  
Db 517 cgttaacaaatctggtctctgaacttaagcaagaacaaacttaacaggtccatccag 576  
Oy 563 atatcttgggaatttgcgtgac---cccgacatatatcttcgcataaccagctca 619  
Db 577 agtcatcttggttcgttccagaaacagtcctgaacttcgcctatcacataaccagctct 636  
Oy 620 ccggttctgtcccaaaacttctgtagagcagatccaattagctgactcctcaggga 679  
Db 637 ccggtccaatcccaataacatcattagcaacacgcagacttaaccggtatcattatccgca 696  
Oy 680 acagactagaagtgatattcattctctgttggcctaaacacgcttgaatgctag 739  
Db 697 acaagctccaagtgatctcgtatgtgttgcctcaacaaacacccggtcattgt 756  
Oy 740 attttcagaagaagtgcttgcattcattctccaggggtgcaggaattccaccctct 799  
Db 757 acttaacagaacatcttccagttcgtatctcaaaagt---tgatatccctaaacac 813  
Oy 800 tgaatacttgacttgaaacataacagtcagcggaagtcgtcgtgagtgactgcta 859  
Db 814 ttggtatcttggaacttgaaacacacatggtgactcgtggaatatccgggtgcagtgactg 873  
Oy 860 aatggaacctgcagacatttaacgctcagtgataataatctctgcgcgaagatccacaag 919  
Db 874 aagctcctctcagttcttcaatgttagctacaacaacagtggtgcacatcccaactg 933  
Oy 920 gggggaacctccagagatttcgaccgctacgctctacccacaacagttgctgtgtg 979  
Db 934 gagggaacctcagaacatttgatcttattcttcaatttccacaacagtggtgtgtg 993  
Oy 980 ctccatgcagaaatgctag 999  
Db 994 ctctctctgaatttgaag 1013

## RESULT 4

AAC33037  
ID AAC33037 standard; DNA; 1132 bp.

XX AC AAC33037;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 1554.

XX KW Hybridisation assay; genetic mapping; gene expression control;

XX KM protein identification; signal transduction pathway;

XX OS metabolic pathway; promoter; termination sequence; ss.

XX PN Arabidopsis thaliana.  
EP1033405-A2.

XX PD 06-SEP-2000.  
XX PE 25-FEB-2000; 2000EP-0301439.  
XX PR 25-FEB-1999; 99US-0121825.  
XX PR 05-MAR-1999; 99US-0123180.  
XX PR 09-MAR-1999; 99US-0123548.  
XX PR 23-MAR-1999; 99US-0125788.  
XX PR 25-MAR-1999; 99US-0126264.  
XX PR 29-MAR-1999; 99US-0126785.  
XX PR 01-APR-1999; 99US-0127462.  
XX PR 06-APR-1999; 99US-0128234.  
XX PR 08-APR-1999; 99US-0128714.  
XX PR 16-APR-1999; 99US-0129845.  
XX PR 19-APR-1999; 99US-0130077.  
XX PR 21-APR-1999; 99US-0130449.  
XX PR 23-APR-1999; 99US-0130510.  
XX PR 23-APR-1999; 99US-0130891.  
XX PR 28-APR-1999; 99US-0131449.  
XX PR 30-APR-1999; 99US-0132048.  
XX PR 30-APR-1999; 99US-0132407.  
XX PR 04-MAY-1999; 99US-0132484.  
XX PR 05-MAY-1999; 99US-0132485.  
XX PR 06-MAY-1999; 99US-0132486.  
XX PR 06-MAY-1999; 99US-0132487.  
XX PR 07-MAY-1999; 99US-0132863.  
XX PR 11-MAY-1999; 99US-0134256.  
XX PR 14-MAY-1999; 99US-0134218.  
XX PR 14-MAY-1999; 99US-0134219.  
XX PR 14-MAY-1999; 99US-0134221.  
XX PR 14-MAY-1999; 99US-0134370.  
XX PR 18-MAY-1999; 99US-0134768.  
XX PR 18-MAY-1999; 99US-0134941.  
XX PR 20-MAY-1999; 99US-0135124.  
XX PR 21-MAY-1999; 99US-0135353.  
XX PR 24-MAY-1999; 99US-0135629.  
XX PR 25-MAY-1999; 99US-0136021.  
XX PR 27-MAY-1999; 99US-0136392.  
XX PR 28-MAY-1999; 99US-0136782.  
XX PR 01-JUN-1999; 99US-0137222.  
XX PR 03-JUN-1999; 99US-0137528.  
XX PR 04-JUN-1999; 99US-0137502.  
XX PR 07-JUN-1999; 99US-0137724.  
XX PR 08-JUN-1999; 99US-0138094.  
XX PR 10-JUN-1999; 99US-0138540.  
XX PR 10-JUN-1999; 99US-0138847.  
XX PR 14-JUN-1999; 99US-0139119.  
XX PR 16-JUN-1999; 99US-0139452.  
XX PR 17-JUN-1999; 99US-0139453.  
XX PR 17-JUN-1999; 99US-0139492.  
XX PR 18-JUN-1999; 99US-0139454.  
XX PR 18-JUN-1999; 99US-0139455.  
XX PR 18-JUN-1999; 99US-0139456.  
XX PR 18-JUN-1999; 99US-0139457.  
XX PR 18-JUN-1999; 99US-0139458.  
XX PR 18-JUN-1999; 99US-0139459.  
XX PR 18-JUN-1999; 99US-0139460.  
XX PR 18-JUN-1999; 99US-0139461.  
XX PR 18-JUN-1999; 99US-0139462.  
XX PR 18-JUN-1999; 99US-0139463.  
XX PR 18-JUN-1999; 99US-0139750.  
XX PR 18-JUN-1999; 99US-0139763.  
XX PR 21-JUN-1999; 99US-0139817.  
XX PR 22-JUN-1999; 99US-0139899.  
XX PR 23-JUN-1999; 99US-0140353.  
XX PR 23-JUN-1999; 99US-0140354.  
XX PR 24-JUN-1999; 99US-0140695.  
XX PR 28-JUN-1999; 99US-0140823.  
XX PR 29-JUN-1999; 99US-0140991.  
XX PR 30-JUN-1999; 99US-0141287.  
XX PR 01-JUL-1999; 99US-0141842.  
XX PR 01-JUL-1999; 99US-0142154.







PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145222.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145293.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147302.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149829.  
PR 23-AUG-1999; 99US-0149802.  
PR 25-AUG-1999; 99US-0149930.  
PR 26-AUG-1999; 99US-0150566.  
PR 27-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158222.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 25.1%; Score 250.8; DB 21; Length 1164;  
Best Local Similarity 55.7%; Pred. No. 4; 2e-71;  
Matches 343; Conservative 0; Mismatches 422; Indels 9; Gaps 3;

QY 32 tggcatatgcatgatttctctctgctcccaactctctgcatcaagaatgcaaca 91  
DB 46 tgtcttgcctctgccaactctctctcccaagactctctgataagaatctctgata 105  
QY 92 acaagcagaagcaagcttacttcccaataacaaacagccttgaaacccacattacag 151  
DB 106 aagatgacaaactaccctccctcaagaatcaagaatcccttaacacacttaccactcg 165  
QY 152 actcatgggtgctcagaacgagatgtgtgttggagcctagtcgaatgtgaagaacca 211  
DB 166 cctcatgggatcccaaaacagactgtgtcttgttactgctcgtgctgctgctgctgca 225  
QY 212 g---caaccgataattccctcataatccaagaacgagacgtctccagggccaatcc 268  
DB 226 ctgtcaaccacgctgtcaacttccctcaatacaatacaagaacgagacgtctgctgcatcc 285  
QY 269 caactcaggtggagacactacatacctccaagccttactgtgtcgttcaatccccaactc 328  
DB 286 cccctgaagtcgtactacgatacgtatcttaaccccttactctccgaaactactaac 345  
QY 329 ttttcgaaaaatccccaagaagaatttctgcaactcaagaactaaatccctcagactca 388  
DB 346 tcaactggtacatccaacccaactatcgcaagctcaagaactcaacttcttcagactaa 405  
QY 389 gctcgaccagtcctcagtgccctgtcccttattcttccctcagcttcgaactaact 448  
DB 406 gctggaaccaatcctcagcgtccgctcgttctcgtgattcttgagtcagctcaagaatcttga 465  
QY 449 gtttagactatcgtttaacaacttttgggtgaatcccccagcttcactccttc 508  
DB 466 acattgaccttctcttaactatgactctctcgttccatacactagcttccctccctac 525  
QY 509 cgaaccttaagccctcgaacttagaagctaaagcaactcaccggtgaaatcccgatatt 568  
DB 526 gtaaacctcgatcttgaacttagtagaacaagcttaacaggtccgataccagagcat 585  
QY 569 ttgggaatttgcctgga---tcccggaacatatacttctgcataaccagctcaggt 625  
DB 586 ttgggaacttccaggaaggtccctagcttcttccatacacaacagactctctgta 645  
QY 626 tcttcccaaaacttctgtagagcagatcaacttaagcctcagcttcaggaaagagc 685  
DB 646 ctataccaaggtcactagcgaatcccgacttaccgcatcgattatccgaaacagc 705  
QY 686 tagaagtgatatttcaatcttcttgggcttaaaaaacgcttgaatgtcagatattt 745  
DB 706 ttaagcgatgtcttaacttcttctgtagcttaaaaaacaaatgactgcttgaat 765  
QY 746 caggaacgtgcttagttcaacttctccagggtgagaggttccacccctttagaat 805

Db 766 caagaaatgttcacgttcgtctcccaagt---taagctgctaagacactaata 822  
Oy 806 acttagacttgaccataaccagatcagcgaagctgtcgaagtaattggaattgg 865  
Db 823 acttgacatgaatacattgacatcacaaggagatattccgctcgaagcgaagctt 882  
Oy 866 accctgcacacatttaacgttcagtgataataatctcgcgcgaagattccaacaggaggaa 925  
Db 883 atttcagtaactgagttagctataacagacgtgtggaagcatccccaaggagagt 942  
Oy 926 accctcagaagttgcagcgtcagcgtctctccacaacagttgctgtgtgctccat 985  
Db 943 atattcaagaatttgattcttattcttctccacaacagtgattatgtgtgaccttc 1002  
Oy 986 tgcagaatgctag 999  
Db 1003 ttctcagtgcaag 1016

RESULT 6  
AAC48837  
ID AAC48837 standard; DNA; 1157 BP.  
XX  
AC AAC48837;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
De Arabidopsis thaliana DNA fragment SEQ ID NO: 58956.  
XX  
KM Hybridisation assay; genetic mapping; gene expression control;  
KM protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139494.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140685.  
PR 28-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140981.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.



Db 1000 ttccctagtcgaag 1013

RESULT 7

AC AAC39926 standard; DNA; 1175 BP.

AC AAC39926;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 26411.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

PN EPI033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123380.

PR 09-MAR-1999; 99US-0123348.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 07-MAY-1999; 99US-0132486.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 21-MAY-1999; 99US-0135124.

PR 24-MAY-1999; 99US-0135353.

PR 25-MAY-1999; 99US-0135629.

PR 27-MAY-1999; 99US-0136021.

PR 28-MAY-1999; 99US-0136392.

PR 01-JUN-1999; 99US-0136782.

PR 03-JUN-1999; 99US-0137222.

PR 04-JUN-1999; 99US-0137528.

PR 07-JUN-1999; 99US-0137502.

PR 08-JUN-1999; 99US-0137724.

PR 10-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 14-JUN-1999; 99US-0138847.

PR 16-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139453.

PR 18-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 21-JUN-1999; 99US-0139763.

PR 22-JUN-1999; 99US-0139889.

PR 23-JUN-1999; 99US-0140353.

PR 24-JUN-1999; 99US-0140354.

PR 28-JUN-1999; 99US-0140635.

PR 29-JUN-1999; 99US-0140823.

PR 30-JUN-1999; 99US-0140991.

PR 01-JUL-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 02-JUL-1999; 99US-0142154.

PR 06-JUL-1999; 99US-0142055.

PR 08-JUL-1999; 99US-0142390.

PR 12-JUL-1999; 99US-0142920.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 20-JUL-1999; 99US-0144352.

PR 21-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145089.

PR 23-JUL-1999; 99US-0145124.

PR 23-JUL-1999; 99US-0145125.

PR 23-JUL-1999; 99US-0145145.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145911.

PR 27-JUL-1999; 99US-0145918.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 03-AUG-1999; 99US-0146389.

PR 04-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 05-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147312.

PR 06-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 09-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 16-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150586.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158322.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 21.8%; Score 218; DB 21; Length 1175;  
Best Local Similarity 54.0%; Pred. No. 1,9e-60;  
Matches 518; Conservative 0; Mismatches 430; Indels 12; Gaps 3;

QY 42 catgatttctcgtgctcccaactctctgcatcaacaagatgnaaaacagacaa 101  
DB 93 ctctcttccatcccttctcaactctcccatcttcttaagctgcacgaaatgacaa 152  
QY 102 gcaagcttaccatcaacaacagcgttgnaaaacccaccatcaagacatcaggt 161  
DB 153 gaatgcctctctcaaatcaagaagcacttgacatccctctctatccctcggaa 212

QY 162 gtacagacgacgattgtgtgtgttggaacctagtcgaatgtgacgnaaacagacgacat 221  
DB 213 ccccgacacgactgtgtgtacgctgtgacggcgt-----cgatgcactaacgctcg 266  
QY 222 aatttccctataatcaagaagcagcagcgtctccacggcgaacatccaccatagtg 281  
DB 267 cgttaaccggtcctcagcgttaaccctccggaagtcctccggtacatctcgtaacagatcgg 326  
QY 282. agaccctaccatcccaagccttatgtgtcgttaactccccaatctttcggaanaat 341  
DB 327 cgacctgtgtacctcgtctacacttagattcagctactctgcacacctaacccggaanaat 386  
QY 342 cccagagaagaattctgcgactcaagaacctaaatccctcagatcagctgcagcagct 401  
DB 387 cccacgacacatccacaacatcaagaacctcaacactctctatcttgaagcacactagctc 446  
QY 402 cagtgccctgtcccttattcttccctcagctcagcttcaagcactatgttgaactatc 461  
DB 447 ctccggtccgattccctgattacatcagcgagctcaagaagcctcagcttcttgaactctc 506  
QY 462 gttaacaacttttggtgttaactccctccctcagcttccactcttcggaacctaaagc 521  
DB 507 tttaacccaatttacccggtcccaattcccggttgcctatctcagatgcggaagcttgaagc 566  
QY 522 cctgcacttagaagcgaactcaacggttgaatcccgatatactttggaatttgc 581  
DB 567 catccaatcaacgataatagtaacggttctataccaacactcttcggttcttctgt 626  
QY 582 tgg---atcccggaatatactttcgcataacacgctcaccgggttgttcccaaac 638  
DB 627 tggcaacgtccctaaactcctcaactgtctaaacaaagcgttcaggaanaataacagaaac 686  
QY 639 ttctgttagagcagatcccaattagctgcactctcaggaacagacatgaaggtgatat 698  
DB 687 gtgtcttaataacgacttcaacgacggtgattgttcgggaataggttccgaagagatgc 746  
QY 699 ttcattctgttggggtcctaaacacgcttggaaatgctcagatttccaggaacgtgct 758  
DB 747 cttaattgtcttcgcaggaacaaacacagtaacgagttgagctgtcgagaacacgtgt 806  
QY 759 taattcaattctccaggtgtgagaggtttccaccctcttgcatacttgaactgtga 818  
DB 807 taattcgcattcgtcgaagt---taagtcgttagaagacatagttcgttgatctcgag 863  
QY 819 ccaataacagatcagcgaagctgtcgaatgtaattgctaatttgcagctcagacatt 878  
DB 864 ccagaacacatctcaggggaagataccacacgactgacttaactacacctcagacatt 923  
QY 879 taagtcagatgataataactctcgtcggaagatcccaacagggggaacctccagagatt 938  
DB 924 caacgtgagcgaacacacatctcgtcggaanaatacccaagttgcgtctcttcacagact 983  
QY 939 cgacgtacgactatctccacaacaggtgctgtgtgtgtgcacatcgcagaaatgta 998  
DB 984 tgaacatctgccttcgcctcaacaacatcgtcttctgtggaactccctcgaagcttgta 1043

RESULT 8  
AAC36540  
ID AAC36540 standard; DNA; 1079 BP.  
XX AAC36540;  
AC AAC36540;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14181.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
OS Arabidopsis thaliana.  
XX



QY 449 glltgacttalcgtttaacaacaacttgggctgtaatccctccatcgcttccacttcc 5008  
 Db 467 acattgcgcttctccatgaatgacctcttggttccataccagttccctccatcatc 526  
 QY 509 cgaaccttaagacctgacttagaagctaaagcactaccggttgaaatcccgatattc 568  
 Db 527 gtaaatctcgatattcttgaactagtagaacaacgcttaagcgtacacgaatcat 566  
 QY 569 ttgggaattctgtcggaac---cccggaacatactcttcgcabaacacgctaccggt 625  
 Db 587 ctggagcgtttccaggaaaggtccctcgtcttcttcatacacaacacgactctcgtta 646  
 QY 626 ttgttcccaaacacttctgtctagagacagatccaattagcgtccgacttccaggaaacagc 685  
 Db 647 ctatacacaagctcaactagcgaatcccgactttacgcgcatcgattatctcgaaacagc 706  
 QY 686 tagaagctgatatctcaactcttggtttggcgcctaaaacgcttggaaatgcatatttt 745  
 Db 707 ttcaaggcgaatgcttcatcatcttggttgagctaaaaaaacaatgtagcgtttgataat 766  
 QY 746 caggaaacgtgctagttagttaaattctccagggtgcag 783  
 Db 767 caagaaacatgcttccagttcgatcttccaaaggttaag 804

OS	Arbidiopsis	Thaliana.
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000.	2000EP-0301439
XX		
25-FEB-1999.	99US-0121825	
PR	03-MAR-1999.	99US-0123180
PR	09-MAR-1999.	99US-0123348
PR	23-MAR-1999.	99US-0125788
PR	29-MAR-1999.	99US-0126264
PR	01-APR-1999.	99US-0126785
PR	06-APR-1999.	99US-0127462
PR	08-APR-1999.	99US-0128234
PR	16-APR-1999.	99US-0128714
PR	19-APR-1999.	99US-0129845
PR	21-APR-1999.	99US-0130077
PR	23-APR-1999.	99US-0130549
PR	28-APR-1999.	99US-0130510
PR	30-APR-1999.	99US-0130891
PR	30-APR-1999.	99US-0131449
PR	04-MAY-1999.	99US-01320407
PR	05-MAY-1999.	99US-0132484
PR	06-MAY-1999.	99US-0132485
PR	06-MAY-1999.	99US-0132486
PR	07-MAY-1999.	99US-0132487
PR	11-MAY-1999.	99US-0132863
PR	14-MAY-1999.	99US-0134256
PR	14-MAY-1999.	99US-0134218
PR	14-MAY-1999.	99US-0134219
PR	14-MAY-1999.	99US-0134221



PR 14-MAY-1999: 99US-0134370.  
PR 18-MAY-1999: 99US-0134768.  
PR 19-MAY-1999: 99US-0134941.  
PR 20-MAY-1999: 99US-0135124.  
PR 21-MAY-1999: 99US-0135353.  
PR 24-MAY-1999: 99US-0135629.  
PR 25-MAY-1999: 99US-0136021.  
PR 28-MAY-1999: 99US-0136392.  
PR 01-JUN-1999: 99US-0136782.  
PR 03-JUN-1999: 99US-0137222.  
PR 04-JUN-1999: 99US-0137528.  
PR 07-JUN-1999: 99US-0137502.  
PR 08-JUN-1999: 99US-0137724.  
PR 10-JUN-1999: 99US-0138094.  
PR 10-JUN-1999: 99US-0138540.  
PR 14-JUN-1999: 99US-0138647.  
PR 16-JUN-1999: 99US-0139119.  
PR 16-JUN-1999: 99US-0139452.  
PR 17-JUN-1999: 99US-0139453.  
PR 17-JUN-1999: 99US-0139492.  
PR 18-JUN-1999: 99US-0139454.  
PR 18-JUN-1999: 99US-0139455.  
PR 18-JUN-1999: 99US-0139456.  
PR 18-JUN-1999: 99US-0139457.  
PR 18-JUN-1999: 99US-0139458.  
PR 18-JUN-1999: 99US-0139459.  
PR 18-JUN-1999: 99US-0139460.  
PR 18-JUN-1999: 99US-0139461.  
PR 18-JUN-1999: 99US-0139462.  
PR 18-JUN-1999: 99US-0139463.  
PR 18-JUN-1999: 99US-0139750.  
PR 18-JUN-1999: 99US-0139763.  
PR 21-JUN-1999: 99US-0139817.  
PR 22-JUN-1999: 99US-0139899.  
PR 23-JUN-1999: 99US-0140353.  
PR 23-JUN-1999: 99US-0140354.  
PR 24-JUN-1999: 99US-0140695.  
PR 28-JUN-1999: 99US-0140823.  
PR 29-JUN-1999: 99US-0140991.  
PR 30-JUN-1999: 99US-0141287.  
PR 01-JUL-1999: 99US-0141842.  
PR 01-JUL-1999: 99US-0142154.  
PR 02-JUL-1999: 99US-0142055.  
PR 06-JUL-1999: 99US-0142390.  
PR 08-JUL-1999: 99US-0142803.  
PR 09-JUL-1999: 99US-0142920.  
PR 12-JUL-1999: 99US-0142977.  
PR 13-JUL-1999: 99US-0143542.  
PR 14-JUL-1999: 99US-0143624.  
PR 15-JUL-1999: 99US-0144005.  
PR 16-JUL-1999: 99US-0144085.  
PR 16-JUL-1999: 99US-0144086.  
PR 19-JUL-1999: 99US-0144325.  
PR 19-JUL-1999: 99US-0144331.  
PR 19-JUL-1999: 99US-0144332.  
PR 19-JUL-1999: 99US-0144333.  
PR 19-JUL-1999: 99US-0144334.  
PR 19-JUL-1999: 99US-0144335.  
PR 20-JUL-1999: 99US-0144352.  
PR 20-JUL-1999: 99US-0144632.  
PR 21-JUL-1999: 99US-0144884.  
PR 21-JUL-1999: 99US-0144884.  
PR 21-JUL-1999: 99US-0145086.  
PR 21-JUL-1999: 99US-0145088.  
PR 22-JUL-1999: 99US-0145085.  
PR 22-JUL-1999: 99US-0145087.  
PR 22-JUL-1999: 99US-0145089.  
PR 22-JUL-1999: 99US-0145192.  
PR 23-JUL-1999: 99US-0145145.  
PR 23-JUL-1999: 99US-0145218.  
PR 23-JUL-1999: 99US-0145224.  
PR 26-JUL-1999: 99US-0145276.  
PR 27-JUL-1999: 99US-0145913.  
  
PR 27-JUL-1999: 99US-0145918.  
PR 27-JUL-1999: 99US-0145919.  
PR 28-JUL-1999: 99US-0145951.  
PR 02-AUG-1999: 99US-0146386.  
PR 02-AUG-1999: 99US-0146388.  
PR 02-AUG-1999: 99US-0146389.  
PR 03-AUG-1999: 99US-0147038.  
PR 04-AUG-1999: 99US-0147204.  
PR 04-AUG-1999: 99US-0147302.  
PR 05-AUG-1999: 99US-0147192.  
PR 05-AUG-1999: 99US-0147260.  
PR 06-AUG-1999: 99US-0147303.  
PR 06-AUG-1999: 99US-0147416.  
PR 09-AUG-1999: 99US-0147493.  
PR 09-AUG-1999: 99US-0147935.  
PR 10-AUG-1999: 99US-0148171.  
PR 11-AUG-1999: 99US-0148319.  
PR 12-AUG-1999: 99US-0148341.  
PR 13-AUG-1999: 99US-0148565.  
PR 13-AUG-1999: 99US-0148684.  
PR 16-AUG-1999: 99US-0149368.  
PR 17-AUG-1999: 99US-0149175.  
PR 18-AUG-1999: 99US-0149426.  
PR 20-AUG-1999: 99US-0149722.  
PR 20-AUG-1999: 99US-0149723.  
PR 20-AUG-1999: 99US-0149929.  
PR 23-AUG-1999: 99US-0149902.  
PR 23-AUG-1999: 99US-0149930.  
PR 25-AUG-1999: 99US-0150566.  
PR 26-AUG-1999: 99US-0150884.  
PR 27-AUG-1999: 99US-0151065.  
PR 27-AUG-1999: 99US-0151066.  
PR 27-AUG-1999: 99US-0151080.  
PR 30-AUG-1999: 99US-0151303.  
PR 31-AUG-1999: 99US-0151348.  
PR 01-SEP-1999: 99US-0151930.  
PR 07-SEP-1999: 99US-0152263.  
PR 10-SEP-1999: 99US-0153070.  
PR 13-SEP-1999: 99US-0153758.  
PR 15-SEP-1999: 99US-0154018.  
PR 16-SEP-1999: 99US-0154039.  
PR 20-SEP-1999: 99US-0154779.  
PR 22-SEP-1999: 99US-0155139.  
PR 23-SEP-1999: 99US-0155486.  
PR 24-SEP-1999: 99US-0155659.  
PR 28-SEP-1999: 99US-0156458.  
PR 29-SEP-1999: 99US-0156596.  
PR 04-OCT-1999: 99US-0157117.  
PR 05-OCT-1999: 99US-0157753.  
PR 06-OCT-1999: 99US-0157865.  
PR 07-OCT-1999: 99US-0158029.  
PR 08-OCT-1999: 99US-0158232.  
PR 12-OCT-1999: 99US-0158369.  
PR 13-OCT-1999: 99US-0159293.  
PR 13-OCT-1999: 99US-0159294.  
PR 13-OCT-1999: 99US-0159295.  
PR 14-OCT-1999: 99US-0159329.  
PR 14-OCT-1999: 99US-0159330.  
PR 14-OCT-1999: 99US-0159331.  
PR 14-OCT-1999: 99US-0159637.  
PR 14-OCT-1999: 99US-0159638.  
PR 18-OCT-1999: 99US-0159584.  
PR 21-OCT-1999: 99US-0160747.  
PR 21-OCT-1999: 99US-0160767.  
PR 21-OCT-1999: 99US-0160768.  
PR 21-OCT-1999: 99US-0160770.  
PR 21-OCT-1999: 99US-0160814.  
PR 21-OCT-1999: 99US-0160815.  
PR 22-OCT-1999: 99US-0160980.  
PR 22-OCT-1999: 99US-0160981.  
PR 22-OCT-1999: 99US-0160989.  
PR 25-OCT-1999: 99US-0161404.  
PR 25-OCT-1999: 99US-0161405.

PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161997.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162142.

Query Match	18.1%	Score 181:	DB 21:	Length 1518:
Best Local Similarity	52.5%	Pred. No. 2.4e-48:		
Matches 551; Conservative	0;	Mismatches 420;	Indels 78;	Gaps 4;

26 ctttttgatgatgatgatgattttcccttcgtccctccaaaccttcctgcatcaagaagt 85  
1363 GCGAATTGTGTCTTCTTGTCTTGTTCACATTCCTTCACAGCCTGTTTGTCTTAAGATCTCT 1304  
86 gcaacaacaagacagaagaagcttactctcaaatcaaaacaagccttgaanaaaaccacaca 145  
1303 GTAACCAAAATGACAAAAACACCCCTCCAGATCAACAAATCTTAACAAACCCCTTATC 1244  
146 ttcaagactatgggtgtgcagacgaagattgttgggttgggaacctaaagatgtgaag 205  
1243 ACCTCGCTTCATGGGACCCTTAACCCGACTGTGTCTCGTGTACTGCCCTTGATGTGCGGGG 1184  
206 aaacagac-----aacgcataattcccccataatccaagacgaagcctccacagcgc 262  
1183 ACGGCACCGTTAACCACCGGTGTACCGGCTTACCATATTTCTCCGGCCAGATCTCCGGTC 1124  
263 aaatcccaacctcaggttggagacctacatacctccaagccttaagtgtcgtlaaacctc 322  
1123 AGATTCGCGCGTGAAGTCGGTACTGTCCGTATCTTGAAGACCCCTTCTTCCCAACCTCT 1064  
323 ccaatcctttcgggaanaaattccagaagaatttccgcactcaagaagccttaaatccctca 382  
1063 CTAACTCACCGGGTACATCCACCCACCATTCGCCAACTCAAGAACTCCGAATGCTCA 1004  
383 gactcaagctcgacagctcagtcagtcgacctgtcccttattctccctcagcttaagaac 442  
1003 GGCTCAGCTGGACCAATCTGACAGGTCCAAATTCGACTTATAGTACGCTCAAGAAATC 944  
443 taactgttttgacttttcgtttaaaaaactttgggtgtatccctccacagctttcca 502  
943 TCGAGTTCTTAGAATCTTCTTCAATGATCTCTGTGTTCATTTCCAAAGTCTCTCTCTA 884  
503 cctctccgaaccttaagaagccttcacttagaagcgtlaagaacctcaccggtga----- 555  
883 CGTTACCTAAATCTTGGCTCTTGAACTTAGCAGGAACAAACTTACAGTTAAAGAACAT 824  
556 ----- 555  
823 CTYTTCTTATTACATCATTCAGTGTCTTGTCTAAMACGATCTTTATGTTGAGTT 764  
556 --attcccgatactttgggaatttctgtagtc---ccggaatatatactttgcata 610  
763 CCAATACCAGATCTATTGGTGTCTTCCAGGAACAGTCCCTGACCTTGCCCTATACATA 704  
611 accagctcaccgggttctgtcccaaaactttgtctagaagcagatccaaattagagctcgaact 670  
703 ACCAGCTCTCCGGTCCCAATACCCAAATCATATGAGCAACATGACATTTAACGGATCGATT 644  
671 ttctcagaagaagcctgaaggtgatatttcaatctctgtttggcgcttaaaaaagccttg 730  
643 TATCCCGCAACAGCTCCAAAGTGATGCTTGATGTGTGTGTGTTCCAAACAAACACCT 584  
731 aaatgcataattttcagaagaagcgtgtcagttcaattctccagaaggtgcagaagtttc 790  
583 GGTCTATTGACTATTCAAAGAAATGTTCCAGTTCGATTCGATCTCAAAAGT---TGATATCC 527  
791 caaccttctgaactactttagacttgaacctaaacgaatcagaagcgaagctgtgcagtg 850  
526 CTAAACACTGTGTATCTTGACTTGAACCAATGGATGACTGCGAATATTCGGGTGC 467

Oy	851	aatgcgcctaaattggactctgcaagacatttaacgfcagtgataataaactctctgcgcgaaga	910
Db	466	AGTGGACCTGAAGCTCCTCTTACGTTCTTCATGTTTACTACAAACACTGTGTGGACACA	407
Oy	911	ttccaaacgggggaacctccagagattgcgcgtaeagcctatctccacaacagttgct	970
Db	406	TCCCAACTCGGAGGGGAACCTTGACACTTTGATTCTTATTCCTATTTCACAAACAAGTGT	347
Oy	971	tgtgtgtgtcctccatlgccagaatgctag	999
Db	346	TGTGTGTGTCTCTCTTGAAATTTGCAAG	318

RESULT	10
AAC52792/C	
ID	AAC52792 standard; DNA: 1528 BP.
XX	
AC	
XX	AAC52792:
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 72454.
XX	
KW	Hybridisation assay; genetic mapping; gene expression control;
KM	protein identification; signal transduction pathway;
XX	metabolic pathway; promoter; termination sequence; ss.
OS	Arabidopsis thaliana.
XX	
PN	EPI033405-A2.
XX	
PD	06-SEP-2000.
XX	
FE	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	11-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	18-MAY-1999; 99US-0134780.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135553.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136021.
PR	27-MAY-1999; 99US-0136392.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137528.
PR	04-JUN-1999; 99US-0137502.



```

OY 47 ttctctcgtcctccaacccctctgcatacaagaatgcacaacgaagcaag 106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1344 TGTTCACATTCCTCAGACCTGTTGTCTAAAGATCTGTGAACCAATAAGCAAAACA 1285
OY 107 cttactccaatacaaacagccttgaaaaccccaactcaagctcagtggtcag 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1284 CCTCTCAAGATCAAGAAATCTTAACACACCTTATACCTCGCTTATCGGACCTTC 1225
OY 167 acgacgatgtctgtgcttggaactagtcgaatgtgaacaaacagc---aacgcataa 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1224 AAGCGACATGTTGCTCTCTGACTGCTTGAGTGCAGCGCACCCACCGTTAACACCGTG 1165
OY 224 ctccccaataatcaagaagaagctctcacgcggaacaaatccacactcagtggtgag 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1164 TTACCGCCCTTACCAATATCTCCGGCAGATCTCCGGTCAGATCCGGCTGAAAGTCGGTG 1105
OY 284 accaccataactccaagccttaagcttgctcgtfaaactcccaactcttcggaaaaatcc 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1104 ACTTGCCGTTATCTTGAGACCTTGTTCTCCGCAAACTCTTAACCTCACCGGTACATCC 1045
OY 344 cagaagaanaattctgcactcaagaactaaatccctcagactcagctcgaagctca 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1044 AAGCCACATCGCAAACTCAAGAACCTCCGAATGCTCAGGCTCAGCTGAGCAATCTGA 985
OY 404 gtggccctgtcccttaattctccctcagctcagaactaactcgtttagactatcgt 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 984 CAGGTCCAAATCTCTGACTTATAGTCACTCAGATCTCAGATCTTCTTGAACCTTTCTCT 925
OY 464 ttaacaactcttggtggtgaatccctccctcagctcttcacacttcggaactaaagccc 523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 924 TCAATGATCTCTCTGTTTCAATTCAGTTCCTCTCTACGTTACTTAATCTTGCTTC 865
OY 524 tgcacttaagaactgaactcaccgtgaa-----atcccgataltcttgga 555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 864 TTGAAGCTTAGCAGGAACAACCTTACAGGAAGAAGATCTTCTCTTATTAATCAAT 805
OY 556 -----atcccgataltcttgga 574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 804 CAGTCTCTTGTCTAAACAAGATCTTATGTTGTAGTTCATACAGTCAAGTCTTGGGT 745
OY 575 atttgcgtgagatc---cccgagacataatcttcgataacagctcaccggttctc 631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 744 CATTTCAGGAACAGTCCCTGACCTTCGCTTATCAATACAGCTCTCCGGTCCAAATTC 685
OY 632 ccaaaacttctgtagacagataccaattaggctcagactctcaggaagaacgactagaag 691
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 684 CCAATTCATTAAGCAACAATCGACTTAAACCGGATGATTTATCCCGCAACAAGCTCCAG 625
OY 692 gtagataattcatctctgttggtggtcctaaaaacgcttggaatgtcagattttcagaa 751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 624 GTGATGCTTGATGTTGTTGTTGTTCCAAACAAACACCTGCTTATTAATTAACAA 565
OY 752 acgtcgttagttcaattctccaggtgcaggaagttccacccctcttgacatactag 811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 564 ACATGTTCCAGTTCATATCTCAAAAGT---TGATATCCCTTAAACACTTGTGTAATCTGG 508
OY 812 acttgaaccataacagatcagcgaagctcgtcgaagtgaattgtgctaattgtgacctgc 871
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 507 ACTTGAACACACATGGGATCACTGGGAATATTCCTCGTACAGTGAGTCAAGCTCTCTTC 448
OY 872 agacatttaacgtcagtgataataatctctcgcgaagaatcccaaggggaaacccccc 931
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 447 AGTTTTCATGTTAGCTACACACACTGTGTGGACACATCCCACTGAGGGAACCTTC 388
OY 932 agagaatcgaacgtcagcgtatctccacaacagtgctgtgtgtcgtccatgtccag 991
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 387 AGACATTTGATCTTATCTATTTTTCACAAAGAGTGTGTGTGTCTCTCTTGAA 328
OY 992 aatgctag 999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 327 TTTGCAAG 320

```

```

RESULT 11
AA042595
ID AA042595 standard: DNA; 1117 BP.
XX
AC AA042595:
XX
DT 01-OCT-1993 (first entry)
XX
DE Sequence encoding fungine endopolylgalacturonase inhibitor (PGIP).
XX
KM Fungine endopolylgalacturonase inhibitor (PGIP); PCR: primer:
XX pesticide: ss.
XX
OS Phaseolus vulgaris, strain Saxa.
XX
FH Key 1.1026 Location/Qualifiers
FT CDS /*tag= a
FT
XX
PN W09311241-A.
XX
PD 10-JUN-1993.
XX
PE 04-DEC-1992; 92MO-IT00158.
XX
PR 06-DEC-1991; 91IT-0RM0915.
XX
PA (CNDR ) CONSIGLIO NAZ DELLE RICERCHE.
XX (UNGE-) UNIV GEORGIA RES FOUND INC.
PI Albersheim P, Bergmann C, Gervone F, Darvill A;
PI De Lorenzo G, Salvi G;
DR WPI: 1993-197059/24.
XX P-PSDB: AAR37817.
XX
PT Nucleotide sequence contained in recombinant vector - encodes
PT endo-polygalacturonase inhibitor (PGIP), useful for conferring
XX e.g. fungal resistance to plants
XX
PS Claim 3; Page 18-19; 36pp; English.
XX
CC The PGIP is capable of inhibiting activity of the fungine
CC endo-alpha-1,4-D-poly-galacturonase enzyme (PG). PGIP is useful
CC in plant transformation to confer fungal resistance or resistance
CC to microorganisms pathogenicity.
XX
SO Sequence 1117 BP; 287 A; 339 C; 201 G; 290 T; 0 other;

Query Match 17.3%; Score 172.4; DB 14; Length 1117;
Best Local Similarity 51.8%; Pred. No. 1.3e-45;
Matches 526; Conservative 0; Mismatches 466; Indels 24; Gaps 5;

OY 2 tgaatatgaatcatcttctgcgcctatcttggcagatagtgatattctctcgccttc 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11 tcaatatcccgtagaacatcttctcgaagcttaagcataatttgytcaattctgtatct 70
OY 62 caaacctctcgtcatcacaagaatgcacaacgaacgaacgaagcttattccaataa 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 71 tgaagaacgtcctctctgagctatgcaaacccaagaataagaagcccttccataatca 130
OY 122 aaacagccttgaaaaaccacacattacagactcagtggtgtcagacgaagatgtgtg 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 131 agaagaaccttggaacccaacactctctctcatgtgtcacaacacgaatgtgtgta 190
OY 182 gttyggac---ctagtgaatgtgacgaacgaacgaacgaatatttccctataatc 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 191 acagaacctgtcgtgttattgtgacacccgaacccaacacacataatgcgtcaaacacc 250
OY 239 aagagcaggaagctctcaccggccaa-----atcccaactcagtggtgagacc 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	251	tcgacctctccgcccataactcctccaaaacctacacctatccctctccctccgcgaac	310
Qy	287	taccatacctccaaagccttatggtctcgtaaacctcccaatccttttcggaaaatccac	346
Db	311	tccccaataactttctataacttggcgcgtacataaactctgtgcgcgaatccccc	370
Qy	347	aagaaattcttgcactcaaaagacctaanaatccctcagactgaactcgacagttccag	406
Db	371	ccgcatactgcataaactcccaactcccaactctctctatataactacacaaagtctcg	430
Qy	407	gccctgcaccttatcttctccctagcttgaagaaactctgttttagattatctgtta	466
Db	431	ggcgaatacccgattctctcttgcacagatcaaaaccccctgcgaacctgcattctccaa	490
Qy	467	acaaactttgggtgtaatccctccctcagcttccacttccacttcccgaaacctaaagcc	526
Db	491	agccctctccgcgaacctctccctccctccatctctctctccccaactctggcgatca	550
Qy	527	acttgaagcgtaaagaaactcaagcgtgtaaatcccgatatcttcttggaaatttgc---	583
Db	551	cattcgaaaggcaacggaatctccgcgcacatcccgcaactcctaagctgttttcgaagc	610
Qy	584	gatacccgacatatatcttctcgataacagacgcacacgggttgttcccaaatcttg	643
Db	611	tgtttaagcggaatgacatccctccgcgaacccgcctcacagggaagatccacagacgtt	670
Qy	644	ctagagcagatcccaattagctcgcgaactctccaggaacagacagatgaagtattcst	703
Db	671	cgaaatcgaacttgcgttctgttgactgttctctcgaaacatgctcgagggtgtgaagctcg	730
Qy	704	tctgttcttggcctaaaacacgcttggaaatgctagaatttttcaggaaaagttcgtagt	763
Db	731	tgtgtctcgggtctcagataagaacagagaagatatcatcttgcgaagaactctctgtct	790
Qy	764	tcaattctctccagggtgcaggaatttccaccctcttgaatactatacttaagcttgaacca	823
Db	791	tctgattctggggaagc---ggggcttgcataagaactttgaacgggtcttgagatctaga	847
Qy	824	accagatccagcggaagctctgctcagatgaaattgtgcta---atgagacccgcagac	880
Db	848	acgtctctatggaagcctacccctcaggaactaacgacgttaagaattcttcgcaagttaa	907
Qy	881	acgtcagatgataataatctctcgcgcaagatcccaacagggggaacacctccagagatt	940
Db	908	atggaagcttcaacaactctgctgcgtggaagattcctcaagttggaacttgaagaagtt	967
Qy	941	acgtgaagcctatcccaacaagaattgcttctgtgtgtgtcctcatctgcagaatgc	996
Db	968	acgttctctctatgccaacaaacaaagtgtctgtgtgtctctctctctctccctcc	1023

	RESULT	12
	AAT86755	
ID	AAT86755	standard; DNA; 2917 bp.
XX		
AC	AAT86755;	
DT	05-FEB-1998	(first entry)
XX		
DE	Raspberry drul gene promoter-PGIP chimeric gene.	
XX		
KM	Promoter; drul; raspberry; transgenic plant; fruit; ripening;	
KW	tissue-specific gene expression;	
XX	polygalacturonase inhibitor protein; PGIP; ds; ss.	
OS	Chimeric - Rubus idaeus.	
OS	Chimeric - Class Planta.	
XX		
FH	Key	
FT	Promoter	
FT		
CAAT-signal		
/note= "drul gene promoter"		
1106..1109		
/tag= a		
Location/Qualifiers		
16..1362		

FT	TATA_signal	/*tag- a
FT	1267..1273	
FT	/*tag- b	
FT	1305..1310	
FT	/*tag- c	
FT	1366..2394	
FT	/*tag- d	
XX		
XX	WO9727308-A1.	
XX		
XX	31-JUL-1997.	
XX		
XX	27-JAN-1997;	97MO-US01443.
XX		
XX	29-JAN-1996;	96DS-0592936.
XX		
XX	(AGRI-) AGRITOE INC.	
XX		
XX	Bestwick RK, Kellogg JA;	
XX		
DR	WPI: 1997-393694/36.	
DR	P-PSDB: AAW29148.	
PT	Plant fruit, especially raspberry, drup promoter - useful for high	
PT	level, tissue specific expression of heterologous sequences,	
PT	especially to modify ethylene production	
XX		
XX		
PS	Example 7: Page 44-47: 67pp: English.	
XX		
CC	This DNA sequence comprises a chimeric gene formed from the	
CC	raspberry drup gene promoter (see AAT86739) and the pathogenesis-	
CC	related polygalacturonase inhibitor protein (PGIP) gene. The PGIP	
CC	coding sequence and 3' untranslated sequence were obtained by	
CC	PCR amplification (see AAT86753-54). The PCR predicted amino acid	
CC	sequence is given in AAW29147. The drup promoter is used to provide	
CC	high-level, tissue-specific expression of PGIP in transgenic plants.	
XX		
XX		
XX	Sequence 2917 BP: 884 A: 638 C: 506 G: 889 T: 0 other:	

	Query Match	17.1%	Score 170.8	DB 18	Length 2917
	Best Local Similarity	51.7%	Prod. No. 6.9e-45		
	Matches 525	Conservative 0	Mismatches 467	Indels 24	Gaps 5
QY	2	tgaatatgaatcatctcttgcgcctattttgtgatataagcatgatttccctgcctc	61		
Db	1376	tcaatatccccgtaacccttgctcttcaagcttaagcataatttggcatctctgcatct	1435		
QY	62	caaacctctctgcatacacaagaatgcaacaacaagaacgaagacgttactcaaatca	121		
Db	1436	tgaagaactgcactctcagcatgcatatgcaaccacaagaataagaagcccttccaaatca	1495		
QY	122	aaagaagcttgaaaaaaccccccacatttaagactataggtgttcagaagactttgtg	181		
Db	1496	agaaagactcttggcaaacccaacacctctctcttctatgtgtcttccaaaccgacttgyta	1555		
QY	182	gttggagac---ctagctcgaaatgtgacgaacacgaacgcgataaattccctcataatgc	238		
Db	1556	acagaacctgtgctagtggttttctgycgaccgcgaacccaacaacatctgcgtcaacaacc	1615		
QY	239	aagcgcgaggaagcctctacccggccaa-----atcccaactcaggttgggaagc	286		
Db	1616	tcgaactctccggcacaataaccctccaaaaccctacccatctcttcctccctgcgaacc	1675		
QY	287	tacacatctcccaagccttatggttcgtaaaactcccaactcttttcggaanaattccag	346		
Db	1676	tcacctactccaatttctctatacatlttgcgcatcataataacctgtgtgctcaatccccc	1735		
QY	347	aagaatatttgcacttcaaaagactcaataatccctcagactcagctcgacagcttcagtg	406		
Db	1736	cgcgcacgcgttaaacataccaacactccacatctctatcttactcacaacaaatggtctcg	1795		
QY	407	ggccttgcaccttatcttccctccagcttcgaacaaactgttttgaacttaagctta	466		

```

Db      1796  ggcgaatacgcgaattctctgtcacagatacaaacctcgtcacccctgcgaacttccctaca 1855
Oy      467  acaaatcttgggtgaatacctccctcagcttccacttccctcgaacttcaaacctcgc 526
Db      1856  agccctctccgcgaacccctccctccatctctctctccccaacctggagaatca 1915
Oy      527  acttaagacgttaacgaactaccgcgtgtaaatcccgatctcttggaaatttgc--tg 583
Db      1916  catcgcagcaacgcgaatctccgcgcacatcccgactccgactcgcgtcttgcgaagc 1975
Oy      584  gatcccggaacataatcttgcataacacgaactccggttcttcccaaaacttgc 643
Db      1976  tgttaccgagcgaactcccgacacccgcgcacgcgggaagtattccacgcgttgc 2035
Oy      644  ctgagacgagatcaatcttaagctcacttccaggaacagacgactagaagtgatattcat 703
Db      2036  cgaatctgaacctgcgctcgttgcactgtctcgaacatgtcgagaggttgaacgcgtgcg 2095
Oy      704  tcttcttgggccttaaaaaacgcttggaaatgtcgaatttccagaaacgltgctagtt 763
Db      2096  tgttctcgggtcagataagaacagagaatatactcgtgcgaagaactctctctgtctt 2155
Oy      764  tcaatttccacagagtgcaaggaattccacacctcttgcataacttgaacttgaaccata 823
Db      2156  ttagatttggggaagt--ggggttgcagaagaacttgaacgggttgcgtcgtgaagaaca 2212
Oy      824  accagatcagcgaagctcgtcgaagtaattggtctaa--attgacccgcgaacattca 880
Db      2213  accgatactatggagacgtacaccacagagactaacgagcttaagttctcgaagaagttaa 2272
Oy      881  acgtcagtgataataatctctcgcgcaagatcccaacagggggaacacctccagagattcg 940
Db      2273  acgtgagcttcaacaactcgttcggtgagatctcctaaggttggaaacttgaagaagtcttg 2332
Oy      941  accgttcgagcctatcccaacaacagttggttgggttgcctcatgccaagaatgc 996
Db      2333  acgttctctctatgccaacaacagttggttgcctctctctctctctctcctcgcgc 2388

RESULT 13
AAQ42596 3
AAQ42596 standard: cDNA to mRNA: 792 BP.
XX
AC      AAQ42596:
XX
DT      01-OCT-1993 (first entry)
XX
DE      Sequence of variant of fungine endopolylacturonase inhibitor (PGIP).
XX
KW      Fungine endopolylacturonase inhibitor (PGIP); PCR; primer;
XX      pesticide; ss.
XX
OS      Phaseolus vulgaris, strain Pintr.
XX
FH      Key
XX      CDS      Location/Qualifiers
XX
FT      CDS      2..685
XX
FT      CDS      /*tag= a
XX
XX
XX      MO9311241-A.
XX
XX      10-JUN-1993.
XX
XX      04-DEC-1992; 92MO-IT00158.
XX
XX      06-DEC-1991; 91IT-ORM0915.
XX
XX      (CNR ) CONSIGLIO NA2 DELLE RICERCHE.
XX      (UNGE-) UNIV GEORGIA RES FOUND INC.
XX
XX      Albersheim P, Bergmann C, Cervone F, Darvill A;
XX      De Lorenzo G, Salvi G;
XX

```

```

DR      WPI: 1993-197059/24.
DR      P-PSDB: AAR37818.
XX
XX      Nucleotide sequence contained in recombinant vector - encodes
PT      endo-polygalacturonase inhibitor (PGIP), useful for conferring
PT      e.g. fungal resistance to plants
XX
XX      Claim 6: page 21; 36pp; English.
XX
XX      The PGIP is capable of inhibiting activity of the fungine
CC      endo-alpha-1,4-D-poly-galacturonase enzyme (PG). PGIP is useful
CC      in plant transformation to confer fungal resistance or resistance
CC      to microorganisms pathogenicity.
XX
XX      Sequence 792 BP; 206 A; 227 C; 153 G; 206 T; 0 other;
XX
XX
XX      Query Match      13.9%; Score 138.8; DB 14: Length 792;
XX      Best Local Similarity 53.4%; Pred: No. 8,9e-35;
XX      Matches 362; Conservative 0; Mismatches 307; Indels 9; Gaps 3:
XX
Oy      325  aatcttctcggaataatcccaagaagaattctcgaactcaaacgaactaaatccctcaga 384
Db      5  aacctcgtcgtcccaatcccccgcgcacatcgttaactcaaccacatccatctctcat 64
Oy      385  ctcaagctcgaacagctcagtgccctgtcccttattcttccctcagcttaagaacta 444
Db      65  atccccaacccaatgtctccgcgcaataaccgaatttcttgcagaatcaaaacccctc 124
Oy      445  acttcttgaacttctcgttcaacaacttgggtgtatctccctccctcagtttccact 504
Db      125  gtaacctcgaacttcccaacaacgcccctccgcgaacctccctccctccatctctct 184
Oy      505  ctccgaaccttaaaacccctgcacttagaagctaaacgaactcaccggttgaatcccgat 564
Db      185  ctcccaaacctcgtatgaatcaacttcgacgcaacacgaatctccgcgcacatcccgac 244
Oy      565  arccttggaaatttgc---tgatcccccgaacatatacttctgcataacacagctcacc 621
Db      245  tcttaagctcatcttctgaagctgttcacgttcgaatgacatctccgaacacgctctacc 304
Oy      622  ggggttcttcccaaaccttctgtagagagatccaatlaaggctcgaactccagggagac 681
Db      305  ggggaagattccgcgcgaagcttgcgaatctgaacctgcgctcgttgcacttgcgtgaac 364
Oy      682  aactagaagtgatataatctctgttcttggccttaaaaaacgcttggaaatgttagat 741
Db      365  atgcctcaggggtgacgcgctggtgtgttctcgatcagataaagaacagcagaagatacat 424
Oy      742  ttccaagaaacgltcctagttcattctccagggttgcaggagttccacacctcttg 801
Db      425  ctgggaagaactctctctgttcttgaatttgagaagatgggggtgt---caagaacttg 481
Oy      802  acatacttagacttgaacataacagatccaggaagctcgttgcagtgatgtgactaa- 860
Db      482  aacgggttcggaatcgaagaacaacgctatctatggaacgctacgcgaaggacgtcgcag 541
Oy      861  -attggaactcgaacacatllaacgtcagtgataataatctctcgtcgaagaatttcaaca 918
Db      542  cttaagttctgcacagtttaattgttggtctcaacaatctgtgcgtgaggttccctcaa 601
Oy      919  gggggaaacctcgaagatcgcacgctacgctatctccacaacagttgctgtgtgtg 978
Db      602  gttggaaacttgcacaacatttcaagttctcgttactgcatgcacaacaagaatgtgtgtg 661
Oy      979  gctcattgcagaatgc 996
Db      662  tctcctctctcgtcgcgc 679

RESULT 14
AAS01017
ID      AAS01017 standard: cDNA: 1400 BP.

```

XX	AA001017;
AC	18-MAY-2001 (first entry)
XX	
XX	Maize disease resistance enhancing protein ZmLR2-1 cDNA.
DE	
XX	
KW	Maize; plant disease resistance; crop; soybean; sunflower; sorghum;
KW	canola; wheat; alfalfa; cotton; rice; barley; millet; plant cell death;
KW	herbicide resistance; ZmLR2-1; polygalacturonase inhibiting protein;
KW	PGIP; ss.
XX	
OS	Zea mays.
XX	
FT	Key
FT	Location/Qualifiers
FT	5'UTR
FT	1..49
FT	/*tag= a
FT	50..1045
FT	/*tag= b
FT	/product= "ZmLR2-1 protein"
FT	3'UTR
FT	1046..1368
FT	/*tag= c
FT	polya_site
FT	1369
XX	/*tag= d
XX	
PN	WO200118061-A2.
XX	
PD	15-MAR-2001.
XX	
PF	06-SEP-2000; 2000WO-US24403.
XX	
PR	09-SEP-1999; 99US-0152988.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	
PI	Simmons CR.
XX	
DR	WPI: 2001-226742/23.
XX	
DR	P-PSDB; AAU00447.
XX	
PT	Novel isolated maize disease resistance polynucleotide useful for
PT	increasing resistance in a plant to disease, controlling cell death,
PT	and conferring resistance to herbicides
XX	
PS	Claim 1; Page 77-79; 90pp; English.
XX	
XX	
CC	The present sequence encoding for maize ZmLR2-1 protein is 1 of 7
CC	novel disease resistance proteins (AAU00443-AAU00449). The ZmLR2-1
CC	is a leucine-rich repeat (LRR) containing polypeptide related to PGP
CC	(polygalacturonase inhibiting protein). These novel maize disease
CC	resistance polynucleotides and polypeptides are useful for enhancing
CC	disease resistance in crops and transgenic plants including maize,
CC	soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
CC	barley or millet. The level of disease resistance protein is increased
CC	by transforming a plant cell with a recombinant expression cassette
CC	which comprises the disease resistance polynucleotide operably linked to
CC	a promoter, or by culturing the plant cell under plant growing conditions
CC	to produce a regenerated plant, or by inducing expression of the
CC	polynucleotide to modulate the disease resistance protein in a plant.
CC	The polynucleotides encoding the disease resistance proteins are useful
CC	for increasing resistance in a plant to disease, controlling cell death,
CC	and conferring resistance to herbicides. They are useful as probes or
CC	amplification primers in the detection, quantification, or isolation
CC	of gene transcripts. They can be used for recombinant expression of
CC	their encoded polypeptides, as immunogens in the preparation and/or
CC	screening of antibodies, and in sense or antisense suppression of the
CC	polynucleotide in a host cell, tissue or plant.
XX	
XX	
XX	Sequence 1400 BP: 263 A; 486 C; 406 G; 245 T; 0 other;
XX	

[illegible]





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2002, 19:29:49 ; Search time 44.14 Seconds

(without alignments)  
5125.771 Million cell updates/sec

Title: US-09-308-140-6

Sequence: 1 atgaatatgaatcatctttt.....ctccattgccagatgctag 999

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Capext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA :  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq :  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq :  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq :  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq :  
5: /cgn2\_6/ptodata/2/ina/PCITUS.COMB.seq :  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq :

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	328.8	32.9	1058	1	US-08-238-163-1
2	250.4	29.1	2075	1	US-08-238-163-3
3	172.4	17.3	1116	1	US-08-244-646-14
4	170.8	17.1	2917	1	US-08-592-9368-20
5	140.4	14.1	792	1	US-08-244-646-16
6	140.4	14.1	792	1	US-08-244-646-16
7	75.6	7.6	3541	4	US-09-180-439-5
8	71.8	7.2	3979	4	US-09-180-439-1
9	71.8	7.2	3979	4	US-09-180-439-7
10	68.8	6.9	3573	4	US-09-180-439-2
11	68.8	6.9	3573	4	US-09-180-439-7
12	68.8	6.9	3573	4	US-09-180-439-7
13	55.2	5.5	4104	4	US-09-353-585-4
14	51.2	5.1	5733	4	US-08-473-553A-1
15	46.8	4.7	3921	2	US-08-567-375-1
16	46.8	4.7	3921	2	US-08-567-375-1
17	46.8	4.7	3921	2	US-08-567-375-1
18	46.8	4.7	3921	2	US-08-567-375-1
19	44.8	4.5	6256	2	US-08-475-891A-3
20	44.8	4.5	6256	2	US-08-475-891A-3
21	44.8	4.5	6256	2	US-08-475-891A-3
22	44.8	4.5	6256	2	US-08-475-891A-3
23	42	4.2	2880	2	US-08-666-271-4
24	39.4	3.9	831	2	US-08-567-375-15
25	39.4	3.9	831	2	US-08-567-375-15
26	34.2	3.4	2291	2	US-08-725-736D-1
27	34.2	3.4	2291	3	US-09-162-366B-1

28	34.2	3.4	2291	3	US-09-161-877B-1	Sequence 1, Appli
29	34.2	3.4	3090	4	US-08-945-983-1	Sequence 1, Appli
30	32.4	3.2	6769	1	US-08-480-784-20	Sequence 20, Appli
31	32.4	3.2	6769	1	US-08-483-553-20	Sequence 20, Appli
32	32.4	3.2	6769	1	US-08-487-002-20	Sequence 20, Appli
33	32.4	3.2	6769	1	US-08-483-554B-20	Sequence 20, Appli
34	32.4	3.2	6769	1	US-08-488-011B-20	Sequence 20, Appli
35	32.4	3.2	6769	1	US-08-850-727-20	Sequence 20, Appli
36	32.4	3.2	6769	5	PCT-US95-10202-20	Sequence 20, Appli
37	32.4	3.2	6769	5	PCT-US95-10202-20	Sequence 20, Appli
38	32.4	3.2	6769	5	PCT-US95-10202-20	Sequence 20, Appli
39	31.4	3.1	5521	4	US-08-975-762-48	Sequence 48, Appli
40	31.4	3.1	5521	4	US-09-295-028-48	Sequence 48, Appli
41	31.4	3.1	5521	4	US-09-106-582-48	Sequence 48, Appli
42	31.2	3.1	2625	3	US-08-804-439A-7	Sequence 7, Appli
43	31.2	3.1	2625	3	US-08-720-229-7	Sequence 101, App
44	30.4	3.0	43795	3	US-08-742-185-101	Sequence 5, Appli
45	30.4	3.0	90050	4	US-09-243-041-5	

## ALIGNMENTS

RESULT 1  
US-08-238-163-1  
Sequence 1, Application US/08238163  
Patent No. 5569830  
GENERAL INFORMATION:  
APPLICANT: BENNETT, Alan  
APPLICANT: LABAVITCH, John M.  
APPLICANT: POMEY, Ann  
APPLICANT: STOTZ, Henrik  
TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESS: Townsend and Townsend Hourie and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/238,163  
FILING DATE: 03-MAY-1994  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1058 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 34..1023  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..1058  
OTHER INFORMATION: /standard\_name="Pear PCIP CDNA"  
US-08-238-163-1

Query Match 32.9%; Score 328.8; DB 1; Length 1058;  
 Best Local Similarity .60.2%; Pred. No. 7.4e-96;  
 Matches 600; Conservative 0; Mismatches 387; Indels 9; Gaps 3;

QY 7 atgaatcattcttcgcccatttcttgatcatgcatgatttccctctgctcacaac 66  
 DB 34 ATGGAAGTCAAGTTCTCCCTTCTCTCCCTTAACCTACTCTCTCTCCGTCCTAAC 93  
 QY 67 ctccctgcatcaacaagatgcaacaacaagaacgaacttactccaataacaaca 126  
 DB 94 CCGGCTCTCTCCGTTCTCTGCAACCCGACACACAAAAGTCCCTTACAAATACAAA 153  
 QY 127 gcccgtgaaaaacccacatctacagactcattgggtgtcagaagatgtgtgtgtg 186  
 DB 154 GCTTCCGGCGCCCTCACTGCTTGGCTCATGAAATCAGACACTGCTGCGATTGG 213  
 QY 187 gacctagtcgaatgtgacgaacacgaacacgcgaataattccccaatcaagaagac 246  
 DB 214 TACTGCGTCACTGTGACTCCACCAACCGCATTAATCTCCACCATCTTTGCCGGC 273  
 QY 247 gaagctctcaacggcaacatccacccacggtggagacacacacacacacacacac 306  
 DB 274 CAGG---TGTCAGGCCAAATCCCGCCCTAGTAGAGACATTGCCATACCTTGAAACCTT 330  
 QY 307 tggctccgtaaacctcccaactcttcggaaaaaacccagaagaataattctgcaactaaa 366  
 DB 331 GAATTCATAGCAACCCCAATCTCACTGGCCCAATCCACCCGCAATTCGCAACCTCAAA 390  
 QY 367 gacttaaatccctcagaactcagctcgaactcagctcagctcagctcagctcagctc 426  
 DB 391 GGACTCAAGTCTTCAGGCTCAGTGGACCAACCTTCAGGCTCTGCTGACTTCTCTC 450  
 QY 427 cctcagcttcaagaacacactgtttagactatcgcttcaacaacacttgggtgtacac 486  
 DB 451 AGCCAACCTCAAGAACTCCTCACTTCTGACTCTCTTCAACAACTCACCAGGTCATC 510  
 QY 487 cctccacagcttccactctcgcgaaccttaaacgcttgaactgaacglaacgaactc 546  
 DB 511 CCCAGCTGCTTTCTGAGCTCCCAACCTCGGCGCTCTGCTGAGACCGCAATTAACCTC 570  
 QY 547 accggtgaaatcccgatcattcttgggaatttgcgtgagac---ccggaacataatc 603  
 DB 571 ACAGGTCATATTCGATATGCTTGGGCACTTATGGAACGTTCCAGACTGTATCTC 630  
 QY 604 tcgacataaccagctcaccgggtgtgttcaccaaaacttcttagagagacatccaatagg 663  
 DB 631 TCCGACACACCAAGCTTTCTGATACATTCCAACTTCCTGATGATGACCTTCACCGC 690  
 QY 664 ctgcgaactctcagggaacagactgaagtgatataattcattcttctgttggccttaaaaa 723  
 DB 691 ATGACTTATACAGCAAGCAAGCTGAGAGTGAAGCATCCGTGATATTTGGGCTGAACAG 750  
 QY 724 cgcttggaaatgtagatttctcagaagaacgctgactgaattcattctccagggtgac 783  
 DB 751 ACAACCCAGATTGTGAGCTGCTCAGGAACCTTGTGAATTAATCTGTGCAAGGT---G 807  
 QY 784 gaattccaccccttcttgatataacttgaacatcacacagacagcgaagctgtg 843  
 DB 808 GAGTTTCCGAAACGTTGACTCTGGTGATATACACCAATATAGATCTAGGGAGATATC 867  
 QY 844 tcgaagtaattgtcctaatttgaacctgcagacacattaaagctcagtaataactctgc 903  
 DB 868 CCAAGTAGATTACGCAAGTAAGTAATTCAGTCTGAGCTGAGCTAACAACAGGCTGTGT 927  
 QY 904 ggcgaagatcccaacagggggaacacccacagaagatcgcagcgtcagcctatctccacaac 963  
 DB 928 GGTTCAGATTCTCTGTGGTGAAGTTGCAAGGCTTGACAGCATATTTATTTCCATTAAC 987  
 QY 964 agttgctgtgtgtgtcctatgcagaaatgctag 999  
 DB 988 CGATGCTTGTGGCGTCTCCACTCCCAAGCTGCAAG 1023

RESULT 2  
 US-08-238-163-3  
 ; Sequence 3, Application US/08238163  
 ; Patent No. 5569830  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BENNETT, Alan  
 ; APPLICANT: LABAVITCH, John M.  
 ; APPLICANT: POWELL, Ann  
 ; APPLICANT: STOTZ, Henrik  
 ; TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL  
 ; NUMBER OF SEQUENCES: 24  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Townsend and Townsend Kourie and Crew  
 ; STREET: Stewart Street Tower, One Market Plaza  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: US  
 ; ZIP: 94105-1493  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/238,163  
 ; FILING DATE: 03-MAY-1994  
 ; CLASSIFICATION: 800  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bastian, Kevin L.  
 ; REGISTRATION NUMBER: 34,774  
 ; REFERENCE/DOCKET NUMBER: 2307E-540  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 543-9600  
 ; TELEFAX: (415) 543-5043  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2075 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 421..1401  
 ; US-08-238-163-3

Query Match 29.1%; Score 290.4; DB 1; Length 2075;  
 Best Local Similarity 58.6%; Pred. No. 2.2e-83;  
 Matches 564; Conservative 0; Mismatches 386; Indels 12; Gaps 3;

QY 44 tgaatttcctctgcttccaaacctctcgtgcatcacaaagatgcaacaagaacgaacagc 103  
 DB 446 TTAATTTTCTTCTCTTCTTCTCTCTCACTACATAAGATGATCAATCCGAAAGCAAAA 505  
 QY 104 aagcttactccaatacaaacagccttgaaaaaacccacacatlaagactcagtggtgt 163  
 DB 506 AAGTCTCTTACAAATAAAGAAAGACTTAGGCATCTTACCAATTTAGCTTATGGGATC 565  
 QY 164 csagaacgagatgtgtgtgtggaactcgtgaatgtgacgaacacgaacacgcgacataa 223  
 DB 566 CAACACACATTTGCTTTACTGTGATGATCAATAAATGTCACGGAACCAACCGGATTA 625  
 QY 224 ttccctcaataatcaagaagcgaagctcgaacggccaacatcccaactcagtgaggag 283  
 DB 626 ATGCTCTCAACCGCTTCCAAAGCAA---TATCTCCGGCCAAATTCGCGCAGCGCTGGAG 682  
 QY 284 acctcaaacctcgaagcctatgttccgttaaacctcccaactccttccggaaaaatcc 343  
 DB 683 ACCTTCATATCTCCAAACATTTGGAATTTATCATATGTTACTATATCTCACCGGAACAATTC 742

OY	344	cagaagaatttcgcgaactcaagaacctaaatccctcagacactcagctcgacacgttca	405
Db	743	CACCTGCATTTGCCAAGCTCACAATTTCAAAATGTTAAAGCTCAGCTTCCTACCTAACCTTA	802
OY	404	gtggccctcgtcccttactctccctcagcttaagaacctactgtcttgaactatcgt	463
Db	803	CAGGTCCGATTCCTGAAATTCCTTATGTGTCAGCTGAAAGATTTGAGTGTGCTCGAATTAAT	862
OY	464	ctaaacaactcttgggtgtatccctctcgaacttcgaactctccgaaccttaagccc	523
Db	863	ACAATCATTTATACCGGACAATATCCTCTCTCTCCCTCTCAGCTTCCCAATTTTGCTAGCGA	922
OY	524	tgcacttagaagcgtcaagaacctcagcggtgaatcccccgaactcttgggaattctgcg	583
Db	923	TGTACTTATGATCGTAAACAACTCACCGGAACATTAACGATATGTTTGGAGATTAAAG	982
OY	584	gatcc-----ccggacataatcttctgcataaccagctcaccggttctgttcccaaa	637
Db	983	GACCAAAATATACCGAGATCTCTACCTTTCACACAAAGCTTGACCGGACATGTGCCGGCAT	1041
OY	638	ctttgtctagacgaatccaatlagctctgactcttccagggaaacagactagaagtgata	697
Db	1043	CTTTAGTGATTTGAATTTTTCACAGCTTGATTTCTCCAGGAATAGCTTGAAGGAGATG	1102
OY	698	ttctatctctgttctgggcccataaaacgccttgaatgtcagatcttccagaagaacgtgc	757
Db	1103	TTTGCTTTTGTTCGGGAGAAATTAAGACAGTCAAGTATGATTTATGTGAGAAATTTAT	1166
OY	758	ttagttccaattctccagggtgcaggagttccaccctctttgacataacttagacttga	817
Db	1163	TGGAGTTTGGATATTTTC---GAAATTCGGAGTTTCTGAGAGCTTGATATCATTTGATTTGA	1216
OY	818	accataaccggaatcaggggaagctctgtcgagtgatgtgctaaattgagcccgagact	877
Db	1220	ATCTAATATCAAAATTTTGTGAGCTTACCAACAGATGTAAGAGTGCACATTCGAGTTT	1279
OY	878	ttaacgcacgtgataataactctctgcgcaagaattccacaagggggaaccccgagat	937
Db	1280	TCATGTGAGTTTAAATAGACTTTGTGGACAGATTCCACAAGTGTGAAGCTTGCGAGACT	1333
OY	938	tcgacgcctacgcgcctatctccacaacagttcgtctgtgtgtgtcctcatctgcagaatgct	997
Db	1340	TTGATATTTACTCTTATTTTCATTAACAATGCTTGTGGCTCTCCCTTCCGGAATGTA	1399
OY	998	ag 999	
Db	1400	AG 1401	

RESULT 3

US-08-244-646-14

Sequence 14, Application US/08244646

Patent No. 5744692

GENERAL INFORMATION:

APPLICANT: Cervone, Felice

APPLICANT: De Lorenzo, Giulia

APPLICANT: Salvi, Giovanni

APPLICANT: Albersheim, Peter

APPLICANT: Darvall, Alan

APPLICANT: Bergmann, Carl

TITLE OF INVENTION: Nucleotide Sequences Coding An

TITLE OF INVENTION: Endopolysialacturonase Inhibitor

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sally A. Sullivan

STREET: 5370 Manhattan Circle Suite 201

CITY: Boulder

STATE: CO

COUNTRY: US

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

```

7 COMPUTER: IBM PC compatible
8 OPERATING SYSTEM: PC-DOS/MS-DOS
9 SOFTWARE: PatentIn Release #1.0, Version #1.25
10 CURRENT APPLICATION DATA:
11 APPLICATION NUMBER: US/08/244,646
12 FILING DATE: 06-JUN-1994
13 CLASSIFICATION: 435
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: IT RM 91A 000915
16 FILING DATE: 06-DEC-1991
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: WO PCT/IT/00158
19 FILING DATE: 04-DEC-1992
20 ATTORNEY/AGENT INFORMATION:
21 NAME: Sullivan, Sally A.
22 REGISTRATION NUMBER: 32,064
23 REFERENCE/DOCKET NUMBER: 19-94
24 TELECOMMUNICATION INFORMATION:
25 TELEPHONE: (303)499-8080
26 TELEFAX: (303)499-8089
27 INFORMATION FOR SEQ ID NO: 14:
28 SEQUENCE CHARACTERISTICS:
29 LENGTH: 1116 base pairs
30 TYPE: nucleic acid
31 STRANDEDNESS: double
32 TOPOLOGY: linear
33 MOLECULE TYPE: DNA (genomic)
34 HYPOTHETICAL: NO
35 ANTI-SENSE: NO
36 ORIGINAL SOURCE:
37 ORGANISM: Phaseolus vulgaris
38 STRAIN: Saxa
39 IMMEDIATE SOURCE:
40 CLONE: lambda PGIP-3.3
41 FEATURE:
42 NAME/KEY: CDS
43 LOCATION: 1..1026
44 FEATURE:
45 NAME/KEY: 3'UTR
46 LOCATION: 1027..1116
47 US-08-244-646-14

```

```

Db 371 CCGCATCCCTTAACCTACCCACCACTCATATCTATATCTACTACACCAATGTCTCCG 430
Qy 407 gccctgccccttatctatctccctcagaactcgaacttaacttgattagatcattcgtta 466
Db 431 GCGCATACCGGATTTCTTGTGTACAGATCAAAACCTGTCACACCTCGACTTCTCTTACA 490
Qy 467 acaaaccttgggtgtaacccctcctcagcttccacttccgaacttaagaagccctgc 526
Db 491 ACGCCCTCTCCGGCACCCCTCCCTCCCTCATCTCTCTCCCAACCTCGAGGAATCA 550
Qy 527 acttagaagtaagagaccacacggggaatcccgatctcttgggaatttgc---tg 583
Db 551 CATTCAGCGGACGCAACGAAATCTCCGGCCATCCCGACTCTCAAGGCTGTTTCAAC 610
Qy 584 gatcccgagacatalatcttcgataaacagctcagcggttggttcccaaaacttgg 643
Db 611 TGTATTGCGGCGATGACATCTCCCGCACCGGCTCACGGGAAGATTCCACGAGCTTGG 670
Qy 644 ctgagacagatccaacttagctgactctcagggagaacagactagaagtgatattcat 703
Db 671 CGAATCTGAACCTGCGGCTTGTGACTTGTCTCGAACAATGCTGGAGGCTGACGCTCGG 730
Qy 704 tctgtgttggtcctaaaaacgcttggaatgctagatttccaggaacggtgtagt 763
Db 731 TGTGTTTCGGGTGAGTAAGAACAACGAAAGATACATCTGGCGAAGAACTCTTGTCTT 790
Qy 764 tcaatttcccaagggtgcaggttccacccctcttcgataacttagactgaacacata 823
Db 791 TTGATTTGGGGAAAGT---GGGGTTGTCAAGAACTTGAACGGGTGGATCTGAGAGACA 847
Qy 824 accagacacggaagctctgcagtgaaatggtctaa---attggaactcagacattca 880
Db 848 ACGGTATCTATGAAACGCTACCTCAGGAGCTTAACGACGCTAAAGTTTTCGAAGTTTAA 907
Qy 881 agctcgtatataatctctcgcggaagaattccaaacaggggaacccctcagagttcg 940
Db 908 ATGTGACTTCAACAATCTGTGGGTGAGATCTCTCAAGCTGGGAAGTTGAAGGTTTGG 967
Qy 941 accgtacgacctatctcccaacacagttgctgtgtgtgtcctcattgcagaatgc 996
Db 968 ACGTTTCTTCTTATGCGCAACAACTGCTGTGTGTCTCTCTCTCTCTCTCTCTCTCT 1023

RESULT 4
US-08-592-936B-20
; Sequence 20, Application US/08592936B
; Patent No. 5783393
; GENERAL INFORMATION:
; APPLICANT: Kellogg, JILL A.
; TITLE OF INVENTION: PLANT TISSUE / STAGE SPECIFIC PROMOTERS FOR
; TITLE OF INVENTION: REGULATED EXPRESSION OF TRANSGENES IN PLANTS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Denlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,936B
; FILING DATE: 29-JAN-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Evans, Susan T.
; REGISTRATION NUMBER: 38,443
; REFERENCE/DOCKET NUMBER: 4257-0012

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2917 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: sequence of the drul:pgip chimeric
; INDIVIDUAL ISOLATE: gene
; US-08-592-936B-20

Query Match 17.1% Score 170.8; DB 1: Length 2917;
Best Local Similarity 51.7%; Pred. No. 6.6e-45;
Matches 525; Conservative 0; Mismatches 467; Indels 24; Gaps 5;

Qy 2 tgaatatgaatcatctcttcgacctatttggctatgacatgatttccctgccttc 61
Db 1376 TCAATATCCCACTAATCATGTCTTAAGCTTAATATTGTCATCTTGTATCTT 1435
Qy 62 caaacctctctgcatcaacaagatgcaacaacagacagacagcttactcaatca 121
Db 1436 TCAGAACTGCACTCTCAGAGCTATGCAACCCCAAGATTAAGCAACCCCTTCAAAATCA 1495
Qy 122 aaacgacctgaaaaaaccacacattacagactcagtggtgctcagaagcagatttgg 181
Db 1496 ACGAAACCTTGCGCAACCAACCTCTCTTTCATGGCTTCAACCAACCGACTTGTGA 1555
Qy 182 gttggagc---ctagtgaatgtgacgaacacagacagacataattccctcaatttc 238
Db 1556 ACAGAACTGGCTTAGGTGTTTATGCGACACCGACACCAACATATGCGCTCAACAC 1615
Qy 239 aagaagacgaagctctcaccgagcaaa-----atcccaactaagtgaggagac 286
Db 1616 TCGACCTCTCCGGCCATTAACCTCCCAAAACCTTACCTATCTCTCTCTCTCTCTCT 1675
Qy 287 taccaactcctcaagccttatggttcgtaaatcccaactcttggaaaaaacccag 346
Db 1676 TCCCTTACCTCAATTTCTATCATTTGGCGGATCAATATCTCTGCTCCAAATCCCC 1735
Qy 347 aagaattctgcactcaagaacctaataatccctcagactcagctcagacagctcagtg 406
Db 1736 CCGCATGCTTAATCTACACCACTCCTATCTATATCTACACCACTATGCTCCG 1795
Qy 407 gccctgccccttatctcctcctcagcttacaagaacttaacttgattagatcattcgtta 466
Db 1796 GCGCAATACCGGATTTCTTGTGTACAGATCAAAACCTGTCACCTCGACTTCTCTACA 1855
Qy 467 acaaaccttgggtgtaacccctcctcagcttccacttccgaacttaagaagccctgc 526
Db 1856 ACGCCCTCTCCGGCACCCCTCCCTCCCTCATCTCTCTCTCCCAACCTCGAGGAATCA 1915
Qy 527 acttagaagtaagagaccacggggaatcccgatctcttgggaatttgc---tg 583
Db 1916 CATTCAGCGGACGCAACGAAATCTCCGGCCATCCCGACTCTCAAGGCTGTTTGAAGC 1975
Qy 584 gatcccgagacatalatcttcgataaacagctcagcggttggttcccaaaacttgg 643
Db 1976 TGTTAACGGCGATGACCAATCTCCCGCAACCGCTCACCGGGAAGATTCCACGAGTTG 2035
Qy 644 ctgagacagatccaacttagctgactctcagggagaacagactagaagtgatattcat 703
Db 2036 CGAATCTGAACCTGCGGCTTGTGACTTGTCTCGAACAATGCTGGAGGCTGACGCTGG 2095
Qy 704 tctgtgttggtcctaaaaacgcttggaatgctagatttccaggaacggtgtagt 763
Db 2096 TGTGTTTCGGGTGAGTAAGAACAACGAAAGATACATCTGGCGAAGAACTCTTGTCTT 2155

```

QY 764 tcaatttccagaggtgcagaggttccacccctttagacaacttagactgaaccata 823  
 Db 2156 TGGATTGGGGAAAGT--GGGGTTGTCAAGAACCTTGAACCGGTTGATCTGAGGAAACA 2212  
 QY 824 accagatcagcggaagctcgtcgaagtgaattgctaa---attgaccctgcyacattta 880  
 Db 2213 ACCGATATCTATGGGACGCTACCTCAAGGACTAAGCAGCTAAAGTTTCTGCAAAAGTTTAA 2272  
 QY 881 acgtcagtgataataatctctcgtcggcgaagatccacaacaggggaaacccccaagatctcg 940  
 Db 2273 ATGTGAGCTTCAACATCTGTGCGGTGAGATTCCTCAAGTGGGAACCTGGAAGTTTG 2332  
 QY 941 accgtagcggcctatctccacaacagttgctgtgtgtgtccattgcagaatgc 996  
 Db 2333 ACGTTTCTTCTTATGCCAACACAGTGTGTGTGTCTCTCTCTCTCTCTCTCTCTC 2388

## RESULT 5

US-09-111-573-20

Sequence 20, Application US/09111573

Patent No. 5929302

GENERAL INFORMATION:

APPLICANT: Kelloff, Jill A.

APPLICANT: Bestwick, Richard K.

TITLE OF INVENTION: PLANT TISSUE / STAGE SPECIFIC PROMOTERS FOR

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/111,573

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/592,936

FILING DATE: 29-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Evans, Susan T.

REGISTRATION NUMBER: 38,443

REFERENCE/DOCKET NUMBER: 4257-0012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 324-0960

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 2917 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: sequence of the drul:PCIP chimeric

US-09-111-573-20

Query Match 17.1%; Score 170.8; DB 2; Length 2917;  
 Best Local Similarity 51.7%; Pred No. 6,6e-45;  
 Matches 525; Conservative 0; Mismatches 467; Indels 24; Gaps 5;

QY 2 tgaatatgaatcatccttctgcctcatttgglycalatgcalgaaatttccctcctc 61  
 Db 1376 TCAATATCCCAAGTAACCATGCTCTTAAGCTTAAGCTAATTTGGTCATCTGTGATCTT 1435  
 QY 62 caaacctctctcaltcacaaagatgaaacaacaagacagaagcttactccaatca 121  
 Db 1436 TGAGAACTGCACTCTCAGGCTATGCAACCCACAAATATGAGAGGCCCTTCTCCAAATCA 1495  
 QY 122 aacagccttgaaaaaacccaccattacagactcatgggtgcagagcagatltgtg 181  
 Db 1496 ACAAGACCTTGGCAACCAACCACTCTCTCTCATGCTTCCAAACCGACTGTGTGA 1555  
 QY 182 gtggagc---ctagtcgaatgtgaagaaacagacagccgacataatttcccacaaatc 238  
 Db 1556 ACAGAACCTGGCTAGTGTGTTATGAGACACGACACCAACCAATATGCGTCACAAAC 1615  
 QY 239 aagacgagaaagctctcaccggcaca-----atcccacctaagtgtggagacc 286  
 Db 1616 TCGACCTCTCCGGCATTAACCTCCCAAAACCTACCTATCCCTTCTCCCTCGCCAAAC 1675  
 QY 287 taccaacctccaagccttaigtgtccgtaaacctcccaatcttttggaaataaccag 346  
 Db 1676 TCCCTACTCTCAATTTTCTATATGCGGCGCATATACCTCGTCCGTCATCCCC 1735  
 QY 347 aagaatttctgcactcaaaagacctaataatccctcagactcagctcagctcagtg 406  
 Db 1736 CCGCATGCTTAACATCAACCAACCACTATCTATATGCTACACCAATGCTCTCG 1795  
 QY 407 gccctgccttattcttccctcagcttaacgaactaactgtttagactatcgttta 466  
 Db 1796 GCGCAATACCCGATTTCTTGTACAGATCAAAACCTCGTCACTCGATCTCTCTCA 1855  
 QY 467 acaacttgggtgtaactcctcctcagcttccacactctccagacttgaagccctgc 526  
 Db 1856 ACGCCCTCTCCGGACCCCTCCCTCCCTCATCTCTCTCCCAACCTCGAGGAATCA 1915  
 QY 527 acttgaacgttaacgaactcaccggttgaatcccgatcttcttggaaatttgc---tg 583  
 Db 1916 CATTGACGGCAGCGATATCGCGGCGCATCCCGACTCTCAAGCTCGTTTGAAGC 1975  
 QY 584 gatcccgagacatalatcttcgcataaacagctcaccgggttgttcccaaaacttgc 643  
 Db 1976 TGTTTACGGCGATACCAATCTCCGCAACCCCTCACCGGAGAAATCCACCAAGTTTG 2035  
 QY 644 ctgagcagatcccaattagctgactctcaggaagaagcttgaagatatttcat 703  
 Db 2036 CGAATCTGAACCTGGCGTTCTGTGACTGTCTCGAATCTGTGAGGTCACCGCTCGC 2095  
 QY 704 tctgttctgggcccataaaacgcttggaaatgctagatlttccaggaagctgtttagt 763  
 Db 2096 TGTGTTCGGGTGATGAAGAACGAAAGATACATCTGCGAAGATCTCTTGCTT 2155  
 QY 764 tcaatttctccaggtgcagaggttccacccctcttgacatacttagacttgaacata 823  
 Db 2156 TGTGATTGGGGAAGT--GGGGTTGTCAAGAACCTTGAACGGGTGATGAGGAACA 2212  
 QY 824 accagatcagcggaagctcgtcgaagtgaattgctaa---attgaccctgcyacattta 880  
 Db 2213 ACCGATATCTATGGGACGCTACCTCAAGGACTAAGCAGCTAAAGTTTCTGCAAAAGTTTAA 2272  
 QY 881 acgtcagtgataataatctcgtcggcgaagatccacaacaggggaaacccccaagatctcg 940  
 Db 2273 ATGTGAGCTTCAACATCTGTGCGGTGAGATTCCTCAAGTGGGAACCTGGAAGTTTG 2332  
 QY 941 accgtagcggcctatctccacaacagttgctgtgtgtgtccattgcagaatgc 996  
 Db 2333 ACGTTTCTTCTTATGCCAACACAGTGTGTGTGTCTCTCTCTCTCTCTCTCTCTC 2388

## RESULT 6

US-08-244-646-16

Sequence 16, Application US/08244646

Patent No. 5744692

## GENERAL INFORMATION:

APPLICANT: Cervone, Felice  
 APPLICANT: De Lorenzo, Giulia  
 APPLICANT: Salvi, Giovanni  
 APPLICANT: Albersheim, Peter  
 APPLICANT: Darvill, Alan  
 APPLICANT: Bergmann, Carl  
 TITLE OF INVENTION: Nucleotide Sequences Coding An  
 TITLE OF INVENTION: Endopolygalacturonase Inhibitor  
 NUMBER OF SEQUENCES: 17  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sally A. Sullivan  
 STREET: 5370 Manhattan Circle Suite 201  
 CITY: Boulder  
 STATE: CO  
 COUNTRY: US  
 ZIP: 80303  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/244,646  
 FILING DATE: 06-JUN-1994  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IT RM 91A 000915  
 FILING DATE: 06-DEC-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/IT/00158  
 FILING DATE: 04-DEC-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sullivan, Sally A.  
 REGISTRATION NUMBER: 32,064  
 REFERENCE/DOCKET NUMBER: 19-94  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303)499-8080  
 TELEFAX: (303)499-8089  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 792 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: *Phaseolus vulgaris*  
 STRAIN: Pinto  
 IMMEDIATE SOURCE:  
 CLONE: PAD-1 (DSM NO. 57446926821)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 2..685  
 FEATURE:  
 NAME/KEY: 3'UTR  
 LOCATION: 686..792  
 US-08-244-646-16

Query Match 14.18; Score 140.4; DB 1; Length 792;  
 Best Local Similarity 53.5%; Pred. No. 2e-35;  
 Matches 363; Conservative 0; Mismatches 306; Indels 9; Gaps 3;

QY 325 aatcttctggaataatccagaagaattcttcgaccccaagaagcccaaatccctcaga 384  
 DB 5 AACCTCGTGGTCCATCCCGCCGCTGCTAAACTCAACCCCACTCCACTATCTCTAT 64  
 QY 385 ctacagctcagacagctcagctgagccctgctcccttatcttcctccagctcagaacta 444  
 DB 65 ATCAACCCACCAATGCTCGCGGGCAATACCCGATTCTTGTCACAGATCAAAACCTC 124

QY 445 actgttaagactatcglttaacaactttgggtgtaactccctccagcttccact 504  
 DB 125 GTACACCTCGACTTCTCTACAAACCCCTCTCCGACCCCTACCTCCCTCATCTCT 184  
 QY 505 ctccgaaccttaagccctgcacttagaagtaacgagaccacccggaaatcccgat 564  
 DB 185 CTCCCAACCTCTGTAGATCATCATCGACGCAACCGAATCTCCGGCCATCCCCAC 244  
 QY 565 atcttgggaatttgc---cgaaccggagacatatacttcgcataaccagctcacc 621  
 DB 245 TCTACGGCTCATTTTTCACAGCTGTTCACGTGATGACCATCTCCCGAACCGCTCAC 304  
 QY 622 ggggttgltcccaaaactttgtctagagacagatccaaltgagctcgacttccaaggaa 681  
 DB 305 GGGAAGATTCCGCGCGCTTTCGCAATCTGAACCTGGCGCTTCTGACTTGTCTGAAC 364  
 QY 682 agactagaagtgatattcattctgtgttggtgacctaaagcgttgaaatgtagat 741  
 DB 365 ATGCTGCAAGGTGACGCGCTCGTGTGTTGCGATCAGTAAGAACACCCAGACATACAT 424  
 QY 742 ttctcaggaacgctgactagttcaattctccaggggtcaggaagttccaccctcttg 801  
 DB 425 CTGGCGAAGAACTCTCTGCTTTGATTGGAGAAAGT---GGGCTGTCAAGAACTTG 481  
 QY 802 acataactagacttgaaccataaccagatcagcggaagtcgtgagatgtgctaa- 860  
 DB 482 AACGGTTGATCTGAGCAACACCGTATCTATGAGAGCTACACGACGAGCTGACGAG 541  
 QY 861 --attgacccgcagacatttaacgctcagtgatataatactccgcaggaagattccaca 918  
 DB 542 CTAAAGTTTCTGCACAGTTTAAATGTAGCTTCANCAATCTGTGGGGTGAATTTCTCA 601  
 QY 919 gggggaacccctcagagatcagacgctacccgctactcccaacagttgcttggt 978  
 DB 602 GGTGGGAACCTTGCAACATTTCAAGTTCTGCTTATGCAACACAAAGTCTGTGTGTGT 661  
 QY 979 gctccattgccaagaatgc 996  
 DB 662 TCTCCTCTCTCTGCTGCTG 679

## RESULT 7

US-09-180-439-5  
 : Sequence 5, Application US/09180439  
 : Patent No. 6225532  
 : GENERAL INFORMATION:  
 : APPLICANT: Dixon, Mark S  
 : APPLICANT: Hatixanthis, Kostas  
 : APPLICANT: Jones, David A  
 : APPLICANT: Jones, Jonathan DG  
 : TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof  
 : FILE REFERENCE: 620 - 53  
 : CURRENT APPLICATION NUMBER: US/09/180,439  
 : EARLIER FILING DATE: 1998-12-06  
 : EARLIER APPLICATION NUMBER: PCT/GB97/01249  
 : EARLIER FILING DATE: 1997-05-08  
 : EARLIER APPLICATION NUMBER: GB 9609681.3  
 : EARLIER FILING DATE: 1996-05-09  
 : EARLIER APPLICATION NUMBER: GB 9619924.5  
 : EARLIER FILING DATE: 1996-09-24  
 : NUMBER OF SEQ ID NOS: 10  
 : SOFTWARE: Patentin Ver. 2.0  
 : SEQ ID NO 5  
 : LENGTH: 3541  
 : TYPE: DNA  
 : ORGANISM: *Lycopersicon esculentum*  
 US-09-180-439-5

Query Match 7.6%; Score 75.6; DB 4; Length 3541;  
 Best Local Similarity 48.2%; Pred. No. 2.7e-14;  
 Matches 275; Conservative 0; Mismatches 289; Indels 6; Gaps 2;







US-09-353-585-4

Query Match 6.9%; Score 68.8; DB 4; Length 3573;  
Best Local Similarity 46.6%; Pred. No. 4, 1e-12;  
Matches 291; Conservative 0; Mismatches 327; Indels 6; Gaps 2;

QY 301 gctcttggttcgtaactcccaactcttccgaaataatcccaagaagaattctgcga 360  
DB 1364 GCTTTTGTCTTCTTATGAAATACGCTTGTAGCTGTTCTTCAAGAAATAGTTTAC 1423  
QY 361 ctcaagaccataatccctcagactcagctcgacagctcgaagctcgaagctccctta 420  
DB 1424 CTAAAGTCTTAAAGTCTTGAATTTGAGTGAATGCTTAAATGCTTAAATGCTTCTGCT 1483  
QY 421 tcttcctcagactcagaaactcgtttagactcttgcgttcaactcgttgcgttgc 480  
DB 1484 TCATTCGGGAATTTGACAACTTGTCTAGTTGAAATCTTGAATTAATGAGCTTTCTGCT 1543  
QY 481 gtaacccctcagactcagactcagactcagactcagactcagactcagactcagact 540  
DB 1544 TCTATTCTGAAGAATAGTTAGTACTAGGCTCTTAATGCTTGAATTTGAGTGAAT 1603  
QY 541 gaaccacacggtgaatcccgatatacttgggaatttgcgtggaatcccgagacatata 600  
DB 1604 GCTTTAATGCTTATTCCTTATTCCTGATTCGGAATTTGACAACTTGTAGTGAAT 1663  
QY 601 ctctcgaataaccagctcagcgggttgcgttgcgttgcgttgcgttgcgttgcgttgc 657  
DB 1664 CTGTATTAATTAATCACTTCTTGTGCTTATTCCTGAAAGAAATAGTTACTTAAGTCTCT 1723  
QY 658 attagagctcagactcagcgggttgcgttgcgttgcgttgcgttgcgttgcgttgc 717  
DB 1724 AATGACCTAGGTTGAGGAGAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCT 1783  
QY 718 aaaaacgctggaagactcagactcagactcagactcagactcagactcagactcagact 777  
DB 1784 CTGAAACACTTGTCTTATGTTGATTTTACAAATACAGCTTCTGCTTATTCCTGA 1843  
QY 778 gtcgaagagctcagactcagactcagactcagactcagactcagactcagactcagact 837  
DB 1844 GAATAGGTTAGTACTGAGTCTTCTTACTTACTTACTTACTTACTTACTTACTTACTT 1903  
QY 838 agctcgtcagactcagactcagactcagactcagactcagactcagactcagactcagact 894  
DB 1904 CTATATTCCTGCTTATTCGCAATATGAGAAATCTGCAAGCTCTGATTCATGATGAAT 1963  
QY 895 aatcctcgtcagactcagactcagactcagactcagactcagactcagactcagactcagact 918  
DB 1964 AATCTCATTTGGGAAATTCCTTCA 1987

RESULT 12  
US-09-353-585-1  
Sequence 1, Application US/09353585  
Patent No. 6287865  
GENERAL INFORMATION:  
APPLICANT: Dixon, Mark S  
Jones, Jonathan D  
Jones, David A  
TITLE OF INVENTION: Plant pathogen resistance genes and uses  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye PC  
STREET: 8th Floor, 1100 No. 6287865th Glebe Road  
CITY: Arlington  
STATE: Virginia  
COUNTRY: United States of America  
ZIP: 22201-4714  
COMPUTER TYPE: IBM PC compatible  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/353,585  
FILING DATE: 15-Jul-1999  
CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12Q 1/68  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/930,277  
FILING DATE: 27-Oct-1997  
APPLICATION NUMBER: PCT/GB96/00785  
FILING DATE: 01-Apr-1996  
APPLICATION NUMBER: GB 9506658.5  
FILING DATE: 31-Mar-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ms Mary J Wilson  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 620-69  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6471 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHEICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Tomato  
STRAIN: Cf2  
FEATURE:  
LOCATION: 1754..5012  
NAME/KEY: sig-peptide  
LOCATION: 1677..1753  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-353-585-1

Query Match 6.9%; Score 68.8; DB 4; Length 6471;  
Best Local Similarity 46.6%; Pred. No. 5, 5e-12;  
Matches 291; Conservative 0; Mismatches 327; Indels 6; Gaps 2;

QY 301 gctcttggttcgtaactcccaactcttccgaaataatcccaagaagaattctgcga 360  
DB 3048 GCTTTTGTCTTCTTATGAAATACGCTTGTAGCTGTTCTTCAAGAAATAGTTTAC 3107  
QY 361 ctcaagaccataatccctcagactcagctcagactcagactcagactcagactcagactcagact 420  
DB 3108 CTAAAGTCTTAAATGCTTCTTATGTTGATTTTACAAATACAGCTTCTGCTTATTCCTGA 3167  
QY 421 tcttcctcagactcagaaactcgtttagactcttgcgttcaactcgttgcgttgcgttgc 480  
DB 3168 TCATTCGGGAATTTGACAACTTGTCTAGTTGAAATCTTGAATTAATGAGCTTTCTGCT 3227  
QY 481 gtaacccctcagactcagactcagactcagactcagactcagactcagactcagactcagact 540  
DB 3228 TCTATTCTGAAGAATAGTTAGTACTAGGCTCTTAAATGCTTGAATTTGAGTGAAT 3287  
QY 541 gaaccacacggtgaatcccgatatacttgggaatttgcgtggaatcccgagacatata 600  
DB 3288 GCTTTAATGCTTATTCCTTATTCCTGATTCGGAATTTGAAACAATGTTGAGTGAAT 3347  
QY 601 ctctcgaataaccagctcagcgggttgcgttgcgttgcgttgcgttgcgttgcgttgc 657  
DB 3348 CTGTATTAATTAATCACTTCTTGTGCTTATTCCTGAAAGAAATAGTTACTTAAGTCTCT 3407  
QY 658 attagagctcagactcagcgggttgcgttgcgttgcgttgcgttgcgttgcgttgc 717  
DB 3408 AATGACCTAGGTTGAGGAGAAATGCTTAAATGCTTATTCCTGCTTATTCCTGCTTATTCCTG 3467



Db 4151 TCCTTACCGCAATCGGAACATGACGAGTTTAACTCTCGATCTCTTTCAACGATC 4210  
QY 617 tcacgggttcttcc 632  
Db 4211 TCTCCGGTAGATACC 4226  
RESULT 15  
US-08-567-375-3  
Sequence 3, Application US/08567375  
Patent No. 5952485  
GENERAL INFORMATION:  
APPLICANT: Ronald, Pamela C.  
APPLICANT: Wang, Guo-Liang  
APPLICANT: Song, Wen-Yuang  
APPLICANT: Szabo, Veronique  
TITLE OF INVENTION: Procedures and Materials for Confering  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/567,375  
FILING DATE: 04-DEC-1995  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,645  
FILING DATE: 29-SEP-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/475,891  
FILING DATE: 07-JUN-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/373,375  
FILING DATE: 17-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-058930  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3921 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(1..2676, 3520..3918)  
OTHER INFORMATION: /product= "Xa-21"  
US-08-567-375-3

Query Match 4.7% Score 46.8: DB 2: Length 3921:  
Best Local Similarity 46.6% Pred. No. 5e-05:  
Matches 150: Conservative 0: Mismatches 172: Indels 0: Gaps 0:

QY 313 cgttaaccccaatcttctggaataacccgaagaattctgactcaagaccta 372  
Db 1147 CTGGAATTGATTAAGATCACAGAGACATTCGCAAGATATTGGCAATCTTATGGCTTA 1206

QY 373 aaatccctcagaactcagctcgaccagctcagtgccctgtcccttatctctccctcag 432  
Db 1207 CAACATCTCTATCTCTGCAACAACAATTTGAGAGGGCTCTTCATCATCTGTTGGCGAG 1266  
QY 433 cttacgaaactaactglttagactatcglttaacaaactttgggtgaatccctcc 492  
Db 1267 CTTAAANAATTTAGGCATTTCTACTGCGCTACGAAANAACACTTGAACGGTTGATCCCTTG 1326  
QY 493 cagcttccacitctccgaaccttaagccctgcaactagaacglaacgaactcaccggt 552  
Db 1327 GCCATAGGAATCTTACTGAACCTTAATATCTTACTGTCGGACCAACAATTCAGTGT 1386  
QY 553 gaaatcccgatatactcttgggaatttgcgtgatacccgacataatactctgcataac 612  
Db 1387 TGGATACCATATCACACTCTCAACCTCACAAACTGTGTGATATAGCCCTTCAACTAAT 1446  
QY 613 cagctcaccgggttcttccca 634  
Db 1447 AACCTTAGTGTCCAAATACCA 1468

Search completed: March 13, 2002, 20:14:05  
Job time: 2656 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2002, 18:35:44 ; Search time 1172.24 Seconds  
(without alignments)  
9157.709 Million cell updates/sec

Title: US-09-308-140-6

Perfect score: 999  
Sequence: 1 atgaataatgaaccatctt.....ctccatgcagaatgtag 999

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 segs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:  
1: em\_estfun:\*  
2: em\_estln:\*  
3: em\_estln:\*  
4: em\_estln:\*  
5: em\_estln:\*  
6: em\_estln:\*  
7: em\_estln:\*  
8: em\_estln:\*  
9: em\_estln:\*  
10: em\_estln:\*  
11: em\_estln:\*  
12: em\_estln:\*  
13: em\_estln:\*  
14: em\_estln:\*  
15: em\_estln:\*  
16: em\_estln:\*  
17: em\_estln:\*  
18: em\_estln:\*  
19: em\_estln:\*  
20: em\_estln:\*  
21: em\_estln:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	203.2	20.3	644	11	BG577452 P23 MP12-
2	202	20.2	742	11	BG127791 EST473437
3	190.8	19.1	626	10	AW217580 EST296294
4	188.8	18.9	573	10	AW928759 EST337547
5	188.4	18.9	647	10	AW737407 EST338750
6	187.4	18.8	589	10	AW217650 EST296364
7	186	18.6	584	11	BF270811 GA_Eb000
8	177.4	17.8	548	10	AT730245 BNLGH1646
9	174.4	17.5	533	11	BG132344 EST465236
10	173.4	17.4	546	10	BE459449 EST414741
11	173.4	17.4	781	11	BG587815 EST489590
12	164.4	16.5	640	10	AW776847 EST335912

13	161	16.1	635	11	BF003501
14	160	16.0	587	10	AV542271
15	157.6	15.8	517	10	BE463191
16	153.8	15.4	513	10	AW624693
17	152.8	15.3	663	11	BG441226
18	152.8	15.3	888	11	BG441233
19	152.6	15.3	569	10	A1507686
20	150.4	15.1	517	10	AW398215
21	142.6	14.3	590	10	AW585990
22	140.4	14.1	639	11	BF424032
23	139.6	14.0	557	10	AV549913
24	138.8	13.9	602	11	B1207435
25	138.8	13.9	740	11	B1208875
26	138.2	13.8	515	11	BG598941
27	136.8	13.7	453	10	AW929041
28	135.8	13.6	579	10	AV542007
29	132.2	13.2	483	11	C95367
30	129	12.9	569	10	AV549817
31	122	12.2	635	11	B1405301
32	121.6	12.2	652	11	BF068634
33	120.6	12.1	573	11	BF054100
34	119.4	12.0	416	10	BE022733
35	116.8	11.7	601	11	B1210636
36	113	11.3	826	11	BG582432
37	112.6	11.3	614	11	BE979949
38	111.4	11.2	430	10	AV411355
39	111.2	11.1	427	10	AV411106
40	111.2	11.1	657	10	AW686216
41	110.2	11.0	487	11	BF070747
42	107.2	10.7	596	11	BG405813
43	107	10.7	469	10	A1938190
44	106.6	10.7	511	10	AV535415
45	106.6	10.7	538	10	AV531396

#### ALIGNMENTS

RESULT 1  
BG577452  
LOCUS BG577452 644 bp mRNA EST 05-JUL-2001  
DEFINITION P23 MP12-ADIS 006 Lambda Zap II library Beta vulgaris cDNA similar to polygalacturonase inhibitor precursor, mRNA sequence.  
VERSION BG577452  
KEYWORDS BG577452.1 GI:14602746  
SOURCE EST.  
ORGANISM Beta vulgaris.  
Beta vulgaris.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Caryophyllales; Chenopodiaceae; Beta.

REFERENCE  
AUTHORS Schneider, K., Weisshaar, B., Borchardt, D.C. and Salamini, F.  
TITLE SNP frequency and allelic haplotype structure of Beta vulgaris expressed genes  
JOURNAL Unpublished (2001)  
COMMENT Contact: Schneider K

ADIS unit  
Max-Planck-Institute for Plant Breeding Research  
Carl-von-Linne Weg 10, 50829 Koeln, Germany  
Tel: 00492215062404  
Fax: 00492215062413  
Email: kscheide@piz-koeln.mpg.de  
basis for pginh marker in sugar beet.  
Location/Qualifiers  
1. 644  
/organism="Beta vulgaris"  
/db\_xref="taxon:161934"  
/clone\_lib="MP12-ADIS 006 Lambda Zap II library"  
/dev\_stage="4 week old pot-grown plants"  
/note="Organ: shoot and root; Vector: Bluescript; cDNA was ligated to a 15-mer oligonucleotide adaptor (Stratagene) and cloned into the EcoRI site of the Bluescript plasmid"

#### FEATURES

source



AM217580  
 LOCUS AM217580 626 bp mRNA EST 18-MAY-2001  
 DEFINITION EST296294 tomato flower buds 3-8 mm, Cornell University  
 Lycopersicon esculentum cDNA clone cTOB6M1 similar to  
 polygalacturonase inhibitor protein, mRNA sequence.  
 AM217580.1 GI:6528454  
 VERSION EST.  
 KEYWORDS tomato.  
 SOURCE Lycopersicon esculentum  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 1 (bases 1 to 626)  
 van der Hoeven,R.S., Bezzerides,J.L., Matera,A.L., Holt,I.E., Liang  
 F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Romling  
 C.M., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.  
 Generation of ESTs from tomato flower tissue  
 Unpublished (1999)  
 CONTACT: CUGI  
 CLEMSON UNIVERSITY Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.  
 FEATURES  
 Source Location/Qualifiers  
 1..626  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cTOB6M1"  
 /clone\_lib="tomato flower buds 3-8 mm, Cornell University"  
 /tissue\_type="flower"  
 /dev\_stage="3-8mm buds"  
 /note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
 XhoI; supplier: Tanksley; Flower buds and flowers were  
 taken from greenhouse plants (4-8 wks old, TA496). They  
 were immediately frozen in liquid nitrogen and then  
 size-separated while remaining frozen."  
 BASE COUNT 188 a 158 c 108 g 172 t  
 ORIGIN

Query Match 19.1%; Score 190.8; DB 10; Length 626;  
 Best Local Similarity 60.1%; Pred. No. 6.9e-45;  
 Matches 356; Conservative 0; Mismatches 227; Indels 9; Gaps 2;  
 QY 44 tgaatttcctcgtccctccaaactctctgcacacaaagtacacaaacgaagaacg 103  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 37 TTAATTTCTTCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 96  
 QY 104 aagcttctcccaatacaaaacagcccttgaaaacccacacattacagactcatgtgt 163  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 97 AAGTCCTTCTCAATAAAGAAAGACTTAGGCAATCCCTTACCATTTAGCTTCATGGGATC 156  
 QY 164 cagacagactatgt 223  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 157 CAACACAGATTGCTGTACTGTGTAAGCTGTAATATGACCGGAAACCAACCGGATAA 216  
 QY 224 ttccctcaataatcaagaagaagaagctccacagcgaacacccactcaggctggag 283  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 217 ATGCTCTCAGCGTCTCCAAACCA--TATCTCCGGCCAAATTCGGCAGCCCTCGAG 273  
 QY 284 accac 343  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 274 ACCTTCATATCTGGAACATTTGCAATTTGATGTTACTATCTACACCGACAAATTC 333  
 QY 344 cagaagaattctgcactcaagaacactaaatccctcagactcagctgcagctcga 403  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 334 CACCTGCATTTGCCAGAGCTCAACAATCTCAAAATGTTAAGGCTCAGCTTCACTAACCTTA 393  
 QY 404 gtggccctgtcccttattcttcctcagacttaagaacactcgtttagactatcgt 463

DB 394 CAGGTCCGATCCCTGAATTCCTTAGTCAAGATTTGACGTTGCTCGAGTTGAT 453  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 464 ttaacaactttgggtgaatccctccctcagcttccacacttccgaacctaaagcc 523  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 454 ACAATCAATTTACCGGAACAAATCCCTTCTCCCTCTCAGCTCGAATTTGTAAGCA 513  
 QY 524 tgcacttaagaagtaagcagcaccggtgaatcccgatcttcttggaaatttgcg 583  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 514 TTTACTTAGATGTAACAACCTCACCAGCAATACCGGAATCGTTTGGAGATTAAAG 573  
 QY 584 gatc-----ccggacatatatcttcgcataaccagctcacggtgtgt 629  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 574 GACCAATATACCAAGATCTCTTCTTACACACAGCTTGACCGGACATGT 625

RESULT 4  
 LOCUS AM928759 573 bp mRNA EST 18-MAY-2001  
 DEFINITION EST337547 tomato flower buds 8 mm to pre-anthesis, Cornell  
 University Lycopersicon esculentum cDNA clone CTC03C9 5', mRNA  
 sequence.  
 AM928759  
 VERSION AM928759 GI:8104154  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 1 (bases 1 to 573)  
 van der Hoeven,R.S., Bezzerides,J.L., Matera,A.L., Holt,I.E., Liang  
 F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Romling  
 C.M., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.  
 Generation of ESTs from tomato flower tissue  
 Unpublished (1999)  
 CONTACT: CUGI  
 CLEMSON UNIVERSITY Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.  
 FEATURES  
 Source Location/Qualifiers  
 1..573  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CTC03C9"  
 /clone\_lib="tomato flower buds 8 mm to pre-anthesis,  
 Cornell University"  
 /tissue\_type="flower"  
 /dev\_stage="buds 8mm-to-preanthesis"  
 /note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
 XhoI; supplier: Tanksley; Flower buds and flowers were  
 taken from greenhouse plants (4-8 wks old, TA496). They  
 were immediately frozen in liquid nitrogen and then  
 size-separated while remaining frozen."  
 BASE COUNT 171 a 142 c 97 g 163 t  
 ORIGIN

Query Match 18.9%; Score 188.8; DB 10; Length 573;  
 Best Local Similarity 60.8%; Pred. No. 2.6e-44;  
 Matches 326; Conservative 0; Mismatches 207; Indels 3; Gaps 1;  
 QY 44 tgaatttcctcgtccctccaaactctctgcacacaaagtacacaaacgaagaacg 103  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 37 TTAATTTCTTCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 96  
 QY 104 aagcttaccacaaatcaaaacagcccttgaaaacccacacattacagactcatgtgt 163  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 97 AATCTCTTACAAATTAAGAAAGACTTAGGCAATCTTACCACTTTAGCTTACGGATC 156







Db 475 AGCTCCCTTAAACACTCCCAATCTCGAGGCTTCGATTGGATAGAAACAAGTAACACT 534  
|||||  
Oy 550 ggtgaatccccgatalactltggaatttgcgtg 584  
|||||  
Db 535 GGTACCAATACCAGAACTTTGTGTATGTTCTAG 569  
|||||  
RESULT 8  
AT1730245 548 bp mRNA EST 11-JUN-1999  
LOCUS BNLGH16465 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar  
DEFINITION to (AB013397) polygalacturonase inhibitor [Citrus jambhiri], mRNA  
sequence.  
ACCESSION AT1730245  
VERSION AT1730245.1 GI:5049097  
KEYWORDS EST.  
SOURCE upland cotton.  
ORGANISM Gossypium hirsutum  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.  
REFERENCE 1 (bases 1 to 548)  
AUTHORS Blewitt, M., Matz, E.C., Davy, D.F. and Burr, B.  
TITLE ESTs from developing cotton fiber  
JOURNAL Unpublished (1999)  
COMMENT Contact: Ben Burr  
Biology Department  
Brookhaven National Laboratory  
Upton, NY 11973, USA  
Tel: 516-344-3396  
Fax: 516-344-3407  
Email: burrb@nsl.bnl.gov  
Seq primer: T3 Primer.  
Location/Qualifiers  
FEATURES  
source 1..548  
/organism="Gossypium hirsutum"  
/cultivar="Acala Maxxa"  
/db\_xref="taxon:3635"  
/clone\_lib="Six-day Cotton fiber"  
/tissue\_type="Immature fiber"  
/dev\_stage="Six days post anthesis"  
/lab\_host="XLI-Blue"  
/note="Vector: pBluescript II KS-"  
BASE COUNT 137 a 174 c 82 g 155 t  
ORIGIN  
Query Match 17.8%; Score 177.4; DB 10; Length 548;  
Best Local Similarity 60.5%; Pred. No. 5.5e-41;  
Matches 292; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

Db 363 CTCAGACGCTTGCTGACCTGACCAATCTTCCGCTCAGTCCCTAATTTTTCAGC 422  
|||||  
Oy 430 cagcttaccgaactactltgttagactatcgtttacaactlttgggtgaatccct 489  
|||||  
Db 423 CAAGTTAAGAACTGATTACTTGAGACCTTCAATTAATACCTCTCGGATCTATTCGA 482  
|||||  
Oy 490 cctcagcttccactctccgaacttaagccctgcacttagaagtgtaacgaacc 549  
|||||  
Db 483 AGCTCCCTTCAACACTCCCAATCTCGATGCTTGCATTTGATGAAACAAGTAAC 542  
|||||  
Oy 550 ggt 552  
|||||  
Db 543 GGT 545  
|||||  
RESULT 9  
BG132344 533 bp mRNA EST 31-JAN-2001  
LOCUS BG132344  
DEFINITION EST465236 tomato crown gall Lycopersicon esculentum cDNA clone  
CTOE7C12 5' sequence, mRNA sequence.  
ACCESSION BG132344  
VERSION BG132344.1 GI:12632532  
KEYWORDS EST.  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
REFERENCE 1 (bases 1 to 533)  
AUTHORS van der Hoeven, R., Sun, H., Cho, J., Uterback, T., Hansen, C., Romning  
C. and Tanksley, S.  
TITLE Generation of ESTs from tomato crown gall tissue  
JOURNAL Unpublished (2001)  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.  
Location/Qualifiers  
FEATURES  
source 1..533  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone\_lib="CTOE7C12"  
/clone\_lib="tomato crown gall"  
/tissue\_type="crown gall"  
/dev\_stage="crown galls from full-grown plants (8 wks old)"  
/lab\_host="SOLR"  
/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:  
XhoI. Four wk old greenhouse plants were stab inoculated  
on stem with Agrobacterium tumefaciens C58 (Dr. T.J. Burr,  
Cornell U.). Galls were allowed to develop for another 4  
wks, when gall tissue was frozen in liquid nitrogen."  
BASE COUNT 156 a 136 c 86 g 155 t  
ORIGIN  
Query Match 17.5%; Score 174.4; DB 11; Length 533;  
Best Local Similarity 60.5%; Pred. No. 4.1e-40;  
Matches 305; Conservative 0; Mismatches 196; Indels 3; Gaps 1;

Db 153 CAACACAGATTGCTGTACTAGCTGTAATAATGTACCGGAACCAACCGGATTA 212  
 Oy 224 ttccctcataaataagacgagcagcctcaccggccaatccaccctcagtgaggag 283  
 Db 213 ATGCTCTACCGCTCTTCCAGGCCAA--TATCTCGGCCCAATTCGGGAGCGTGGAG 269  
 Oy 284 accatacactcccaagcctatgltccgtaaacctcccaatccttcggaaaatcc 343  
 Db 270 ACCTTCATATTCGAAACATTGGAATTCATCATGTACTAATCTCACCGGACAAATTC 329  
 Oy 344 cagaagaattctcgcactcaagaacctaaatccctcagactcagctcagactctca 403  
 Db 330 CACCTGCATATTCGAAAGCTCACAAATCTCAAAATGTTAAGGCTCAGCTTCACTAACCTTA 389  
 Oy 404 gtggccctgctcccttattctcctcagctcagctcaagaactcgtttagactatcgt 463  
 Db 390 CAGGTCCGATCCCTGAATTCCTTACTAGCTGAGAGATTTGACGTTGCTGAGTTGAAT 449  
 Oy 464 ttaacaactttgggtgtaatccctcctcagcttccacactcctcagactcaagaacct 523  
 Db 450 ACAATCAATTTACCGGAAACATCCCTTCTCCCTCTCAGCTTCGAAATTTGCTAGCGA 509  
 Oy 524 tgcacttagaagcgtcaagaactca 547  
 Db 510 TGTACTAGATGCTGAACAACTCA 533

RESULT 10  
 BE459449 546 bp mRNA EST 18-MAY-2001  
 LOCUS EST414741 tomato developing/immature green fruit Lycopersicon  
 DEFINITION esculentum cDNA clone CLEM6H18, mRNA sequence.  
 ACCESSION BE459449  
 VERSION BE459449.1 GI:9503751  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:  
 Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:  
 Asteridae: euasterids I: Solanales: Solanaceae: Solanum;  
 Lycopersicon.  
 1 (bases 1 to 546)  
 Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
 Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,  
 Niernm,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley  
 S.D.  
 Generation of ESTs from tomato fruit tissue, immature green  
 Unpublished (2000)  
 Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.  
 Location/Qualifiers  
 1. 546

TITLE Generation of ESTs from tomato fruit tissue, immature green  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.  
 Location/Qualifiers  
 1. 546  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLEM6H18"  
 /clone\_1db="tomato developing/immature green fruit"  
 /tissue\_type="fruit"  
 /dev\_stage="immature green (5-35 days post-anthesis)"  
 /lab\_host="SOLR"  
 /note="Vector: pBluescriptSKmCUnadapt; site\_1: EcoRI;  
 site\_2: XhoI; Fruit was tagged at 5 dpa (0.5 cm) and  
 harvested at 7 day intervals through 35 dpa. Equal masses  
 of tissue from each stage were combined (including seeds  
 and locules) prior to mRNA isolation."

BASE COUNT

156 a 140 c 98 g 152 t

Query Match 17.4%; Score 173.4; DB 10; Length 546;  
 Best Local Similarity 59.9%; Pred No. 8, 1e-40;  
 Matches 329; Conservative 0; Mismatches 211; Indels 9; Gaps 2.

Oy 164 cagaagacatgtgtgtgtgtggaactagtcgaatgtagcagaacccgaacgcataa 223  
 Db 1 CAACACAGATTGCTGTACTAGCTGTAATAATGTACCGGAACCAACCGGATTA 60  
 Oy 224 ttccctcataaataagacgagcagcctcaccggccaatccaccctcagtgaggag 283  
 Db 61 ATGCTCTACCGCTCTTCCAGGCCAA--TATCTCGGCCCAATTCGGGAGCGTGGAG 117  
 Oy 284 accatacactcccaagcctatgltccgtaaacctcccaatccttcggaaaatcc 343  
 Db 118 ACCTTCATATTCGAAACATTGGAATTCATCATGTACTAATCTCACCGGACAAATTC 177  
 Oy 344 cagaagaattctcgcactcaagaacctaaatccctcagactcagctcagactctca 403  
 Db 178 CACCTGCATATTCGAAAGCTCACAAATCTCAAAATGTTAAGGCTCAGCTTCACTAACCTTA 237  
 Oy 404 gtggccctgctcccttattctcctcagctcagctcaagaactcgtttagactatcgt 463  
 Db 238 CAGGTCCGATCCCTGAATTCCTTACTAGCTGAGAGATTTGACGTTGCTGAGTTGAAT 297  
 Oy 464 ttaacaactttgggtgtaatccctcctcagcttccacactcctcagactcaagaacct 523  
 Db 298 ACAATCAATTTACCGGAAACATCCCTTCTCCCTCTCAGCTTCGAAATTTGCTAGCGA 357  
 Oy 524 tgcacttagaagcgtcaagaactcaagaacctcagactccttcggaaattgctg 583  
 Db 358 TGTACTAGATGCTGAACAACTCAACGACCGGAAACATCCGGAATCGTTGGAGATTTAAG 417  
 Oy 584 gac-----cccgacatatactcttcgcataaccagcctcagcggttgcctcaaaa 637  
 Db 418 GACCAATATACGATCTTACTTCCACACACGCTTGACCGGACATGCGGCGAT 477  
 Oy 638 ctcttgctgagcagatccatcaatcagctcagctcctcagctcagcagcagcagcagcagc 697  
 Db 478 CTTAGGATGATTTGAATTTTCCACGCTTGATTTCTCCAGGAATAGCTTGAGAGATG 537  
 Oy 698 ttcatctc 706  
 Db 538 TTTGGTTT 546

RESULT 11  
 BG587815 781 bp mRNA EST 11-APR-2001  
 LOCUS EST489590 MHAM Medicago truncatula/Gloms versiforme mixed EST  
 DEFINITION library cDNA clone pMHAM-5109 5' end, mRNA sequence.  
 ACCESSION BG587815  
 VERSION BG587815.1 GI:13602879  
 KEYWORDS EST.  
 SOURCE Medicago truncatula/Gloms versiforme mixed EST library.  
 ORGANISM Eukaryota: mixed EST libraries.  
 1 (bases 1 to 781)  
 Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.  
 and Fraser,C.M.  
 ESTs from roots of Medicago truncatula after colonization with  
 Gloms versiforme, 2001  
 Unpublished (2001)  
 Contact: Harrison M.J.  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73401  
 Tel: 580-223-5810  
 Fax: 580-221-7380  
 Email: [mj.harrison@noble.org](mailto:mj.harrison@noble.org)  
 Noble EST name: N386888e TIGR sequence name: MTDDK89TK More  
 information is available at: <http://www.medicago.org>  
 Seq primer: Skmod (CTA gAA CTA gAt CC).

TITLE Unpublished (2001)  
 JOURNAL Plant Biology Division  
 COMMENT The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73401  
 Tel: 580-223-5810  
 Fax: 580-221-7380  
 Email: [mj.harrison@noble.org](mailto:mj.harrison@noble.org)  
 Noble EST name: N386888e TIGR sequence name: MTDDK89TK More  
 information is available at: <http://www.medicago.org>  
 Seq primer: Skmod (CTA gAA CTA gAt CC).





DEFINITION AV542271 Arabidopsis thaliana roots Columbia Arabidopsis thaliana  
CDNA clone R2181d08F 3', mRNA sequence.  
ACCESSION AV542271  
VERSION AV542271.1 GI:8704031  
KEYWORDS EST.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 587)  
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.  
A large scale analysis of cDNA in Arabidopsis thaliana: Generation  
of 12,028 non-redundant expressed sequence tags from normalized and  
size-selected cDNA libraries  
DNA Res. 7, 175-180 (2000)  
JOURNAL 20363093  
MEDLINE  
COMMENT Contact: Erika Asamizu  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/  
FEATURES  
Source location/Qualifiers  
1..587  
/organism="Arabidopsis thaliana"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/clone="R2181d08F"  
/clone\_lib="Arabidopsis thaliana roots Columbia"  
/issue\_type="roots"  
/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2:  
XhoI"  
BASE COUNT 166 a 108 c 151 g 162 t  
ORIGIN  
Query Match 16.0%; Score 160; DB 10; Length 587;  
Best Local Similarity 57.0%; Pred. No. 6.9e-36;  
Matches 333; Conservative 0; Mismatches 245; Indels 6; Gaps 2;  
QY 360 actcaagaacccaataacccctcagactcagccagcagctcagtgagccctgctcctt 419  
DB 586 ACTCAAGAACCTCCGATCTCAGGCTCAGCGATCTGACGATCTGACAGGTCTCAATTCCTTA 527  
QY 420 atctctccctcagctcagcaaacactgcttagactatcgtttaacaaactttggg 479  
DB 526 CTTTAACTCAGCTCAAGAAATCTGAGTTCTAGAACTTCTCAATGATCTCTCTG 467  
QY 480 tgttaacctctcagcttccacacttccgaaacttaagccctgagacttgaacgtta 539  
DB 466 TTCATTCCCAAGTCTCTCTCTACGTACTTAATCTTGCTTGAACCTAGCAGAA 407  
QY 540 cgaactcagcgtgaaatcccgatactcttggaatttgcctgacatc--ccgacat 596  
DB 406 CAAACTTACAGGTCCATATACAGAGCTATTGGTCTGTTCCAGAGACAGTCCCTGACT 347  
QY 597 atactcttcgcaaacagctcaccgggttctcccaaacactttgctagagcagatcc 656  
DB 346 TCGCTATACATTAACCACTCTCCGGTCCCAATACCAATCATTAAGCAACATGACTT 287  
QY 657 aattagcctcagcttctcaggaagacagactagaagtgatatattcattcttctgggac 716  
DB 286 TAAACGGATCGATTATCCCGCAACAGCTCCAGGTGATGCTTGATGTTGTTGTTG 227  
QY 717 taaaaaacgtcggaaatgtagatttccaggaagcgttagttagtattcattctccag 776  
DB 226 CAACAACAAACCTGCTATATGACTTATCAAGAAACATGTTCCAGTTGATATCTCAAA 167  
QY 777 ggtcgaagagttccacccctcttgacacttagacttgaaccataacagacagcagcg 836  
DB 166 GGT---GATATCCCTTAATAACACTTGATGTTGACTTGACCTGAACCAACATGGATCAGTGG 110  
QY 837 aagctcgtcgaatgaaatggtcctaatttgacctgcagacatttaacgtcagtgataataa 896

DB 109 GAATATTCGGGTGACAGTGGAGCTGACTCTCTCTTCAAGTTCTTACGTACACAA 50  
QY 897 tctctgcgcaagattccacagcaggggaacctccagatcgc 940  
DB 49 ACTGTGTGACACATCCCACTGAGGAGGAACCTTCAGACACTTG 6  
RESULT 15  
LOCUS BE463191  
DEFINITION BE463191 517 bp mRNA EST 18-MAY-2001  
BE354414 tomato flower buds 8 mm to pre-anthesis; Cornell  
University Lycopersicon esculentum cDNA clone CTC12c9, mRNA  
sequence.  
ACCESSION BE463191  
VERSION BE463191  
KEYWORDS BE463191.1 GI:9508962  
SOURCE EST.  
ORGANISM tomato.  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
1 (bases 1 to 517)  
van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E., Liang  
F., Upton, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Konning  
C.M., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
Generation of ESTs from tomato flower tissue  
Unpublished (1999)  
COMMENT Contact: Cui  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: http://www.genome.clemson.edu/orders/index.html  
5 prime sequence.  
FEATURES  
Source location/Qualifiers  
1..517  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CTC12c9"  
/clone\_lib="tomato flower buds 8 mm to pre-anthesis,  
Cornell University"  
/issue\_type="flower"  
/dev\_stage="buds 8mm-to-preanthesis"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; Supplier: Tanksley; Flower buds and flowers were  
taken from greenhouse plants (4-8 wks old, TA496). They  
were immediately frozen in liquid nitrogen and then  
size-separated while remaining frozen."  
BASE COUNT 151 a 131 c 85 g 150 t  
ORIGIN  
Query Match 15.8%; Score 157.6; DB 10; Length 517;  
Best Local Similarity 59.3%; Pred. No. 3.4e-35;  
Matches 287; Conservative 0; Mismatches 194; Indels 3; Gaps 1;  
QY 44 tgaattccctcgcctccaacccctcgtcacacaagaatgacaacaacagcaagc 103  
DB 37 TTATTTTCTTTGCTTCTCTCTCTCTCATATAGTATGCAATCCGAAGCAAAA 96  
QY 104 aagcttaccccaataaacaagccttgaaaacccacacattacagactcatggtgtc 163  
DB 97 AAGTCTCTCTCAAAATAAGAAAGACTTAGCAATCTTACCATTTACTTTCATGGATC 156  
QY 164 cagaagcagatgt 223  
DB 157 CAACAACAGATGCTGTACTGTACTGTGTAAATGTGACCGGAAACCAACCGGATAA 216  
QY 224 ttccctcaatattcaagcagcagagctccacgcgcgaatccacactcaagtgagga 283  
DB 217 ATGCTCTACCGCTTCTCAAGCAA---TATCTCCGGCAATTTCCGGCAACCGCTCGAG 273

```
OY 284 accatacactcccaagcctatgtgtccgtaaacctcccaatcttttcggaanaatcc 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 274 ACCTTCCATATCTCGAACAATTGGAAATTTCATCATGTACTAATCTCACCGSACAATTC 333
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 344 cagaagaatcttcgcaactcaagaactlaaaatccctcagactcagctcagcagctca 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 334 CACCTGCAATTCGGAAGCTCACCAATCTCAAAATGTTAAGCTCAGCTTCACTAACCTTA 393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 404 gtggcccgctcccttatcttcctcagactaagaaactaactgtttaagactatcgt 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 394 CAGGTCCGATCCCTGAATTCCTTAGTCAAGATTGACGTTGCTCGAGTTGAATT 453
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 464 ttaacaacttttggtgtaalcctcctcagcttccaccttccgaaacctlaagccc 523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 454 ACAATCAATTTACCGAACAAATCCCTTCTCCCTCTCAGCTTCGAAATTGCTAGCGA 513
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 524 tgca 527
    |||
DB 514 TGTA 517
```

Search completed: March 13, 2002, 19:50:47  
Job time: 4503 sec



```

FT Modified-site 291..293
FT /label="N-glycosylation_site
FT /note="potential"
FT Peptide 25..49
FT /note="corresponds to CNBr peptide fragment which
FT was sequenced directly; the Cys residue at
FT position 27 could not be identified by
FT sequencing"
FT Peptide 215..246
FT /note="corresponds to CNBr peptide fragment which
FT was sequenced directly; the Ser residue at
FT position 232 and the Asn residue at
FT position 238 could not be identified by
FT sequencing"
PN US5569830-A.
XX 29-OCT-1996.
XX
XX 03-MAY-1994; 94US-0238163.
XX
XX 03-MAY-1994; 94US-0238163.
XX
XX 03-MAY-1994; 94US-0238163.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Bennett A, Labavitch JM, Powell A, Stotz H;
XX WPI: 1996-496968/49.
XX DR N-PSDB: AAT49434.
XX
XX DNA construct, comprising poly:galacturonase inhibitor protein DNA
XX useful to confer resistance to fungal, partic. Botrytis cinerea,
XX infection in plant, pref. tomato or strawberry
XX
XX Claim 25; Fig 1; 33pp; English.
XX
XX The pear polygalacturonase inhibitor protein cDNA was obtained using
XX mRNA isolated from mature green fruit and the present amino acid
XX sequence was deduced from it. Pear PGIP exhibits differential
XX inhibition of polygalacturonases from different fungal species;
XX introduction of the pear PGIP cDNA into other plant species alters the
XX susceptibility of these plants to the fungal pathogens against which
XX pear PGIP is effective. For example, transgenic tomato fruit
XX expressing high levels of pear PGIP were shown to be more resistant
XX to the fungus Botrytis cinerea than control fruit.
XX
XX Sequence 330 AA:
XX
XX Query Match 50.28; Score 881.5; DB 17; Length 330;
XX Best Local Similarity 54.7%; Pred. No. 3.2e-82;
XX Matches 181; Conservative 41; Mismatches 106; Indels 3; Gaps 3;
XX
XX 3 IESSEFCILICMFLCLPLNLSASORCQNNNDKALLQIKTKALKPRTTDSWSDDDCCGW 62
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 1 melkfscflstlflssvlnpalsdlcnppdkkvlqikafgpyvlaawksdcdw 60
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 63 DLVECDETSNRIISLIQDDDEALTGOIPQVGDLPYLOALFRRKLPMLFGKIPPEISALK 122
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 61 yctcdstlnrlnstltaag-vsgqpalvagdpyletlefhvntlgtipalaktk 119
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 123 DLKSLRSLSSLSGPPVLPFPOLTKLTCLOLSEFKLLGAVIPPOLSTLPNKLALHLENEL 182
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 120 gllsvllfclaspsslvrcnpkdkvllqikdignpyhlaavdpntdcyvvllcdtx 179
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 183 TGEIPOLFEFGNFGAS-PDIYLSHNOITGFPVKTFARADPRLDFSGNLEGGDISFLRPPK 241
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 180 tgnlpisfgfigvwpdiylshmqisnptstgqmfstidstmklegdaavtqlnk 239
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 242 RLEMLDFSGNVLVFNFSRVOEPFPPSLTYLDLNHNQISGSLSELAKLDTOTFVNSDNLG 301
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 240 tctqvdlsrnllefnlskv-efpstsclsvdlnnkikyslpvettqtnfqlnvsvyrtlc 298

```

```

OY 302 GKIPTCNLQRFDRATYLRHNSCLCGAPLPEC 332
DB 299 gqipvggklsgfdeysyfhmrclcgaplpssc 329
XX
XX RESULT 2
XX AAM09096
XX ID AAM09096 standard; Protein: 327 AA.
XX
XX AAM09096;
XX AC
XX XX
XX 05-MAR-1997 (first entry)
XX
XX Tomato polygalacturonase inhibitor protein.
XX
XX PGIP: polygalacturonase inhibitor protein; fungal resistance;
XX Botrytis cinerea; transgenic plant; tomato; strawberry.
XX
XX Lycopersicon esculentum.
XX
XX US5569830-A.
XX
XX 29-OCT-1996.
XX
XX 03-MAY-1994; 94US-0238163.
XX
XX 03-MAY-1994; 94US-0238163.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Bennett A, Labavitch JM, Powell A, Stotz H;
XX WPI: 1996-496968/49.
XX DR N-PSDB: AAT49435.
XX
XX DNA construct, comprising poly:galacturonase inhibitor protein DNA
XX useful to confer resistance to fungal, partic. Botrytis cinerea,
XX infection in plant, pref. tomato or strawberry
XX
XX Claim 26; Fig 2; 33pp; English.
XX
XX A fragment of tomato genomic DNA without introns and coding for
XX the polygalacturonase inhibitor protein was isolated from mature
XX green fruit. The DNA coding for PGIP can be introduced into of
XX plant species, e.g. strawberry, to alter the susceptibility of
XX these plants to the fungal pathogens against which tomato PGIP is
XX effective. The present PGIP amino acid sequence was deduced from
XX the genomic DNA; it includes a block of 10 tandem leucine-rich
XX repeats with an average length of 24 amino acids, as well as seven
XX potential N-glycosylation sites.
XX
XX Sequence 327 AA:
XX
XX Query Match 48.18; Score 844; DB 17; Length 327;
XX Best Local Similarity 52.8%; Pred. No. 2.3e-78;
XX Matches 171; Conservative 46; Mismatches 103; Indels 4; Gaps 3;
XX
XX 11 LCICMFLCLPLNLSASORCQNNNDKALLQIKTKALKPRTTDSWSDDDCCGWLVCDER 70
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 5 lllvvlflclaspsslvrcnpkdkvllqikdignpyhlaavdpntdcyvvllcdtx 64
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 71 SNRIISLIQDDDEALTGOIPQVGDLPYLOALFRRKLPMLFGKIPPEISALKLKSRLS 130
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 65 tnrlnlvtfygan-issqipaavagdpyletlefhvntlgtipalaktlnlkmrls 123
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 131 STLSGPPVLPFPOLTKLTCLOLSEFKLLGAVIPPOLSTLPNKLALHLENELTGEIPOLF 190
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 124 fnltgplpelfsqlnklillelnyngftgclpsslsqipnllamyldrnklgtlpsf 183
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 191 GNFAGS--PDIYLSHNOITGFPVKTFARADPRLDFSGNLEGGDISFLRPPKRLMLDF 248
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 184 grtkgpnlpdiylshmslgnvpsatgdlntfclstsrnklegdvsftlgnktsqvldl 243

```





PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	14-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0156458.
PR	28-SEP-1999;	99US-0156569.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0158029.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158233.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159328.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Accession	Gene	Protein	Protein ID	Protein Name	Protein Description
Db	7	1c1lfflfttclskdclcnqndkrtllkikkslnmphyhlsawdpqtccswylcsgda	66		
Qy	70	TSNRILSIIODEALNGQIPQVGDLPYLOALMFRKLPNLFGKIPETISALKDKLSURL	129		
Db	67	tvnhvtvcltfsqisqgdpaeavgdpyletlvfrkksnltgtlqplakikhlmlrl	126		
Qy	130	SSTSISGAVPLFFPQDITLCGLDSFNKLLGVIPPOQSTLPLNKAHLHERNELTGEIPDI	189		
Db	127	swntcltqypdflsikhlefelsfnlsgsipsstlcklialelstrnkltgsipes	186		
Qy	190	FGNFAGS-PDIYLSHNOITGCVPKTFARADPIRLDFSGNRLGDISFLFGPKRLEMLDF	248		
Db	187	fgsfqgtvpdlrtilsmnqisgdpkxlglnfdnridlstrnkltqgsasmlfgsnkltwsidl	246		
Qy	249	SGNVLSFNFRRKQEPPEPLTYLDLNHNOISGSLSELAKLDLOTFNVSNNLCKRIPFGG	308		
Db	247	srnmqfistky-dipkklglidlnhngltgnlpyvqteadpifgnvsynklcghiptgg	305		
Qy	309	NLQRPDRATYLNHNSCLCGAPLPEC	332		
Db	306	klqtdfsyfnkclcgapleic	329		
RESULT	4				
AAAG6870					
ID	AAAG6870	standard; Protein: 337 AA.			
AC	AAAG6870;				
XX					
DT	18-OCT-2000	(first entry)			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 59012.				
XX					
KW	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter termination sequence.				
XX	Arabidopsis thaliana.				
OS	Arabidopsis thaliana.				
PN	EP1033405-A2.				
XX					
PD	06-SEP-2000.				
XX					
PF	25-FEB-2000; 2000EP-0301439.				
XX					
PR	25-FEB-1999; 99US-0121825.				
PR	05-MAR-1999; 99US-0123180.				
PR	09-MAR-1999; 99US-0123548.				
PR	23-MAR-1999; 99US-0125788.				
PR	25-MAR-1999; 99US-0126264.				
PR	29-MAR-1999; 99US-0126785.				
PR	01-APR-1999; 99US-0127462.				
PR	06-APR-1999; 99US-0128234.				
PR	08-APR-1999; 99US-0128714.				
PR	16-APR-1999; 99US-0129845.				
PR	19-APR-1999; 99US-0130077.				
PR	21-APR-1999; 99US-0130449.				
PR	23-APR-1999; 99US-0130510.				
PR	23-APR-1999; 99US-0130891.				
PR	28-APR-1999; 99US-0131449.				
PR	30-APR-1999; 99US-0132048.				
PR	30-APR-1999; 99US-0132407.				
PR	04-MAY-1999; 99US-0132464.				
PR	05-MAY-1999; 99US-0132485.				
PR	06-MAY-1999; 99US-0132486.				
PR	06-MAY-1999; 99US-0132487.				
PR	07-MAY-1999; 99US-0132863.				
PR	11-MAY-1999; 99US-0134285.				
PR	14-MAY-1999; 99US-0134218.				
PR	14-MAY-1999; 99US-0134219.				
PR	14-MAY-1999; 99US-0134221.				
PR	14-MAY-1999; 99US-0134370.				
PR	18-MAY-1999; 99US-0134768.				

PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136783.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139110.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147302.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157865.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
Query Match 45.1%; Score 792.5; DB 21; Length 337;  
Best Local Similarity 50.0%; Pred. No. 4, 9e-73;  
Matches 162; Conservative 44; Mismatches 115; Indels 3; Gaps 3;  
OY 11 LCICIFLCLPMLSSORCNNDKALLQITAKNPITTDSSWDDCCGADLVECD-DE 69  
DB 14 ICILILFLILCLISKIDCNQDKLILIKKSLINPYHLASWDPQTCOSWYCLECGDA 73  
OY 70 TSNRIISLIIQDEALGQIPQVGDLPYOLALWFRKLPNLGKRPREISALKDKLSRL 129  
DB 74 LVNHVLTALITSGISQIPAEVGDIPYLELIVFRKLSNLGTLQPTAKLKNLMLTL 133  
OY 130 SSTLSGVPLEFPQTLTKLCIDLSFNKLLGVIPPOLSTLPNLKALHLERMELTGEIPDI 189  
DB 134 SWNLITGPIDTISQILKHLFELSINDLSGSPISLSTLPILALELSRNKLTSIPES 193  
OY 190 FGNFPGS-PDIYLSHNOITGFPKTFARADPRILDFSGNRLGEGDISFLFGKKRLEMDF 248  
DB 194 IGSIFGPVDPDLIRLNQISGPIKSGIDINRILDISRNKLQGDASMLFGSNKLTWSIDL 253  
OY 249 SGWVLSFNFSPRQEPFPLTLTLDLHNQISGLSSESLAKLDLQTFNVSDNNLCGRPTGG 308  
DB 254 SRMFGFISKV-DIPKTLGLIDLNHGLTGNIPQWCEARQGFIVSNYKLCGHIPVGG 312  
OY 309 NLQRPDRATYALHNSCLCAPRPEC 332  
DB 313 KLITFDSYSYFHKCLCGAPLEIC 336  
RESULT 5  
AAG07166  
ID AAG07166 standard; Protein; 326 AA.  
XX AAG07166;  
DT 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4215.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
OS  
XX  
PN EPI033405-A2.  
PD 06-SEP-2000. 1  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 18-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145224.  
PR 23-JUL-1999; 99US-0145228.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160981.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 44.68; Score 784; DB 21; Length 326;  
Best Local Similarity 51.18; Pred. No. 3; Se-72;  
Matches 166; Conservative 45; Mismatches 104; Indels 10; Gaps 6;

QY 16 IFPLCPML-----SASORCNRNDKQALLOIKTALKNPTITDSWVSDDCCGMDLVEC-DE 69  
DB 3 Iffllstllllttslakdchdkdktllkikkslnpyhlaaswpkdcswycieegda 62

QY 70 TSN-RIISLIIODEALTGOIPPOVGDLPYLQALMFRRKLNLEFGKIPEDISALDKLSLR 128  
DB 63 tvnhrvtsllldgde-issqdlppevgdlpyltstlffklcnlgnhgdptlaklknltffir 121

QY 129 LSSTLSGVPVLPFPOLTKLTCLDLSFNKLLGVIPOLSLPLPNKALHLERNELTGCIPO 188  
DB 122 lswntltgvpelfsqklnleyldstfnidsipselslsrkleyelsrnkltgpipe 181

QY 189 IFGNFAGS-PDIYLSHQMLGFWPKTEFARADPIRLDPSGRNLEGDIFLFGPKRLEMLD 247  
DB 182 stfgtsgkvpsllfshmqslgtlpsksignpofryridsrnkigqdasilfgakktwivd 241

QY 248 FSGNVLSEFNFSRYOEPPPSLTLYDLNHNQISGSLSESLAKLDLOTFNVSNNLCGKIPTG 307  
DB 242 isrnmgfdlskv-klaktlnldmhnngitgsipaewskayqqlnvsvyrncgripkg 300

QY 308 GNLQREDRTAYLHNSCLCGAPLPBPC 332  
DB 301 eylqridsysffhncicgapplpssc 325

RESULT 6  
AAG46832  
ID AAG46832 standard; Protein; 326 AA.  
XX  
AC AAG46832;  
XX  
DT 18-OCT-2000 (first entry)  
DE  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 58959.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
OS  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PF  
XX 25-FEB-1999; 99US-0121825.  
PR  
XX 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0132566.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139494.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142395.  
PR 06-JUL-1999; 99US-0142399.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 19-JUL-1999; 99US-0144336.  
PR 20-JUL-1999; 99US-0144337.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 21-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 44.6%; Score 784; DB 21; Length 326;  
Best Local Similarity 51.1%; Pred. No. 3.5e-72;

Matches 166; Conservative 45; Mismatches 104; Indels 10; Gaps 6;

OY 16 IFELCPNL-----SASORCNNDKQALLQITALKNPRTTOSWSDDDCCWDVLEC-DE 69  
DB 3 Iffllstllttslakdclchddkctclikkslnpyhlaawpdkccswylecgcda 62  
OY 70 TSN-RIISLITODEALITGOIPQOVGDLPYQALMFRKLPNLFGEISALKDLKSLR 128  
DB 63 tvnhrveslilqgde-issqilpevgdipyltsllfkltlthlqhlptlaklknltflr 121  
OY 129 LSSYSLSGVPVLPFPQTLKTLCDLSFNKLGVIPOLSTLPNKLALHLENNELTGEIPD 188  
DB 122 lswenltqvpelqslqklnleyldstfndlsqslsslsirklleylslsrnkltgripe 181  
OY 189 IFGNFAGS-PIYISHNOJLTFGVPKTFARADPIRLDPSGNRLGEDISLFGPKRLEMLD 247  
DB 182 sfgtfsgkvpstlftshnglsqtlpslsqnpdftfyldstsrnkltgdaasllfyakkttwlvd 241  
OY 248 FSGVNLSEFNFRVOEPFPPSLTYLDLNNHNOISGSLSESLAKLUDLTFVNSDNLGKTIPTG 307  
DB 242 isrmfsgfdlskv-klaktlnldmhnngltsglpaekskayfqllnvsyrnlcgrltpkg 300  
OY 308 GNLORPRTAYLHNSCLGAPLPFC 332  
DB 301 eylqfidsysffhkcicgaplpsc 325

RESULT 7

ID AAG07165 standard; Protein: 330 AA.

XX AC AAG07165;

XX DT -17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 4214.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
OS  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128274.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134281.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 01-JUL-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 19-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147933.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 28-SEP-1999; 99US-0156459.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 44.6%; Score 784; DB 21; Length 330;  
Best Local Similarity 51.1%; Pred. No. 3,6e-72;  
Matches 166; Conservative 45; Mismatches 104; Indels 10; Gaps 6;

QY 16 IFCLPNL-----SASQRCNNNDKQALLQITALKNPITITDSWSDDCCGMDVECD-DE 69  
DB 7 Iffllstllltslaktchdkdckllklkkslnnpkhsldwpcckccswcylecgda 66  
QY 70 TSN-RIISLIIODDLEALGQIPQVGDLPYLOALWFRKLPMFLGKIPPEISALDKLSLR 128  
DB 67 tvnhvvtlllqdgq-issgqippegdipyltsllfrlntlnthgtpktiklnlftlr 125  
QY 129 LSTSLSGVPPLFPQPLKTLCTCLDSFNKLLGVIPPOQISTLPNKLALHLENNELTGEIPD 188  
DB 126 lswntltgpyefslqklnleyldstfndslsgslpsstslrkleylelstrnkltpipe 185  
QY 189 IFGNFAGS-PDIYLSHNOITGFVPEKTPARADPIRLDFSGNLEGDISTLFPPKRLKMLD 247  
DB 186 sfgtltsgkvpstflshnqslgtlpkslgnpdyfrldstrnklsqdasllfgakktwlvd 245  
QY 248 FSGNVLSEFSGVOEFPPSLTYLDLNNHQISGSLSELAKLDELTFNVSDDNLGCKIPTG 307  
DB 246 lstrmfqfidlskv-klaktlnlmlannhgltsglpaewskayfqllnvsynrlcgripkq 304  
QY 308 GNLOFRRTAYLHNSCLGAPLPEC 332  
DB 305 eylqfidsytfhnkclcgaplpse 329





PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149729.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	30-AUG-1999;	99US-0151080.
PR	31-AUG-1999;	99US-0151303.
PR	01-SEP-1999;	99US-0151438.
PR	07-SEP-1999;	99US-0151930.
PR	10-SEP-1999;	99US-0152363.
PR	13-SEP-1999;	99US-0153070.
PR	15-SEP-1999;	99US-0153758.
PR	16-SEP-1999;	99US-0154018.
PR	20-SEP-1999;	99US-0154039.
PR	22-SEP-1999;	99US-0154779.
PR	23-SEP-1999;	99US-0155139.
PR	24-SEP-1999;	99US-0155486.
PR	28-SEP-1999;	99US-0155659.
PR	29-SEP-1999;	99US-0156458.
PR	04-OCT-1999;	99US-0156596.
PR	05-OCT-1999;	99US-0157117.
PR	06-OCT-1999;	99US-0157753.
PR	07-OCT-1999;	99US-0157865.
PR	08-OCT-1999;	99US-0158029.
PR	12-OCT-1999;	99US-0158232.
PR	13-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161932.
PR	28-OCT-1999;	99US-0161933.
PR	29-OCT-1999;	99US-0162142.

[illegible]

Db	126	lswlnlbpvpelfsqlnkleyldisfindisgslpsisslrkleyelsrnlktgpipe	185
Qy	169	IFGNAGS-PIYVLSHNOITGCVPTFEARADPIRLDFSGNRLEGDISFLFGPKRLEMLD	247
Db	166	sfgtsgkvpstflshmqslsgtlpkslgnpdtfyridisrnlkqdasllfgakktwlvd	245
Qy	248	ESGNVLSTNFSRVOEFPPSLTYLIDLNNQISGSLSELAKLDLOTFNVSDDNNLCGIPTG	307
Db	246	lsmmfqgldskv-klaktlnlmldmnhgltgslpaewskayfqllnvsyrrlgrlpkg	304
Qy	308	GNLQREDRATVLIHNSCLGAPLPEC	332
Db	305	eylgridsysifhncklcgaplpsc	329
RESULT 9			
ID	AAc646830	standard; Protein: 337 AA.	
XX			
AC	AAc646830;		
XX			
DT	18-OCT-2000	(first entry)	
XX			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 58957.		
XX			
KW	Protein identification: signal transduction pathway; metabolic pathway;		
KM	hybridisation assay; genetic mapping; gene expression control; promoter		
KX	termination sequence.		
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000;	2000EP-0301439.	
XX			
PR	25-FEB-1999;	99US-0121825.	
PR	05-MAR-1999;	99US-0123160.	
PR	09-MAR-1999;	99US-0123548.	
PR	23-MAR-1999;	99US-0125788.	
PR	29-MAR-1999;	99US-0126264.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130851.	
PR	28-APR-1999;	99US-0131449.	
PR	30-APR-1999;	99US-0132048.	
PR	30-APR-1999;	99US-0132407.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	06-MAY-1999;	99US-0132487.	
PR	07-MAY-1999;	99US-0132863.	
PR	11-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	
PR	14-MAY-1999;	99US-0134219.	
PR	14-MAY-1999;	99US-0134221.	
PR	14-MAY-1999;	99US-0134370.	
PR	18-MAY-1999;	99US-0134768.	
PR	19-MAY-1999;	99US-0134941.	
PR	20-MAY-1999;	99US-0135124.	
PR	21-MAY-1999;	99US-0135353.	
PR	24-MAY-1999;	99US-0135629.	
PR	25-MAY-1999;	99US-0136021.	
PR	27-MAY-1999;	99US-0136392.	
PR	28-MAY-1999;	99US-0136782.	
PR	01-JUN-1999;	99US-0137222.	

PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 26-JUL-1999; 99US-0145313.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149375.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149829.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151067.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154479.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match

44.6%; Score 784; DB 21; Length 337;

Best Local Similarity 51.1%, Pred. No. 3.7e-72;  
Matches 166; Conservative 45; Mismatches 104; Indels 10; Gaps 6;

OY 16 IFCLPML-----SASORCNDKQALLQIKTALKNPTITDSWVSDDCCGMDLVEC-DE 69  
Db 14 lllllscillltslakdchdkdtkllkikkslnpyhasadpkrdcgswclegda 73  
OY 70 TSN-RITSLIODEALTCGIPQVGDLPYLQALWFRKLPNLFCKIPEETISALKDKSLR 128  
Db 74 cvnhvrslllqdg-e-lsagqipevgdlyltsllffkklntlgthqptlakiknltflr 132  
OY 129 LSSTSLSPVLPFPOLTKTKCLDLSNKLGVIPQLSTLPMLKALHLERNLTGEIPD 188  
Db 133 lswtntltgvppeelslkhleyldsfndlsagsispslslrlxleylslsrnkltgpipe 192  
OY 189 IFGNFAGS-PDIYLSHNOQTGVFKTFARADPIRLDFSGNRLEGDISFLFGPKRLEMD 247  
Db 193 sfgtfsgkvpssflshnqtsgrtkelpgdpdfrldsrnkltggsdasilfgakktwivd 252  
OY 248 FSGNVSEFNSRQVQEPPLTYLDLNNHNOISGSLSELAKLDLQTFNVSDDNLCGRIPG 307  
Db 253 lsrnmfqlslkv-klaklmlndmhnltgslpawekayfqllnveynrlcgrlpgk 311  
OY 308 GNLOREFDRTAYLHNSCLCGAPLPEC 332  
Db 312 eylgrfdysffhncclcgaplpsc 336

RESULT 10  
AG07164  
ID AG07164 standard; Protein: 338 AA.

XX AC AG07164;

XX DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 4213.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 29-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0126785.

PR 06-APR-1999; 99US-0127462.

PR 08-APR-1999; 99US-0128234.

PR 16-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0129845.

PR 21-APR-1999; 99US-0130077.

PR 23-APR-1999; 99US-0130449.

PR 28-APR-1999; 99US-0130510.

PR 30-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 07-MAY-1999; 99US-0132486.

PR 11-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134371.

PR 18-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 21-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 25-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 01-JUN-1999; 99US-0136782.

PR 03-JUN-1999; 99US-0137222.

PR 04-JUN-1999; 99US-0137528.

PR 07-JUN-1999; 99US-0137502.

PR 08-JUN-1999; 99US-0137724.

PR 10-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 14-JUN-1999; 99US-0138847.

PR 16-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139453.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139750.

PR 21-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140655.

PR 26-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 21-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

Accession	Species	Gene	Protein	Length	Score	DB	Length	Score	DB
23-JUL-1999	99US-0145224			338	4.68	21	338	4.68	21
26-JUL-1999	99US-0145276			338	4.68	21	338	4.68	21
27-JUL-1999	99US-0145313			338	4.68	21	338	4.68	21
27-JUL-1999	99US-0145318			338	4.68	21	338	4.68	21
28-JUL-1999	99US-0145351			338	4.68	21	338	4.68	21
02-AUG-1999	99US-0146388			338	4.68	21	338	4.68	21
02-AUG-1999	99US-0146388			338	4.68	21	338	4.68	21
03-AUG-1999	99US-0147038			338	4.68	21	338	4.68	21
04-AUG-1999	99US-0147204			338	4.68	21	338	4.68	21
04-AUG-1999	99US-0147302			338	4.68	21	338	4.68	21
05-AUG-1999	99US-0147392			338	4.68	21	338	4.68	21
05-AUG-1999	99US-0147260			338	4.68	21	338	4.68	21
06-AUG-1999	99US-0147303			338	4.68	21	338	4.68	21
06-AUG-1999	99US-0147416			338	4.68	21	338	4.68	21
09-AUG-1999	99US-0147493			338	4.68	21	338	4.68	21
09-AUG-1999	99US-0147935			338	4.68	21	338	4.68	21
10-AUG-1999	99US-0148171			338	4.68	21	338	4.68	21
11-AUG-1999	99US-0148319			338	4.68	21	338	4.68	21
12-AUG-1999	99US-0148341			338	4.68	21	338	4.68	21
13-AUG-1999	99US-0148565			338	4.68	21	338	4.68	21
13-AUG-1999	99US-0148684			338	4.68	21	338	4.68	21
16-AUG-1999	99US-0149368			338	4.68	21	338	4.68	21
17-AUG-1999	99US-0149175			338	4.68	21	338	4.68	21
18-AUG-1999	99US-0149426			338	4.68	21	338	4.68	21
20-AUG-1999	99US-0149722			338	4.68	21	338	4.68	21
20-AUG-1999	99US-0149723			338	4.68	21	338	4.68	21
23-AUG-1999	99US-0149902			338	4.68	21	338	4.68	21
23-AUG-1999	99US-0149930			338	4.68	21	338	4.68	21
25-AUG-1999	99US-0150566			338	4.68	21	338	4.68	21
26-AUG-1999	99US-0150884			338	4.68	21	338	4.68	21
27-AUG-1999	99US-0151065			338	4.68	21	338	4.68	21
27-AUG-1999	99US-0151066			338	4.68	21	338	4.68	21
30-AUG-1999	99US-0151080			338	4.68	21	338	4.68	21
30-AUG-1999	99US-0151303			338	4.68	21	338	4.68	21
31-AUG-1999	99US-0151338			338	4.68	21	338	4.68	21
01-SEP-1999	99US-0151930			338	4.68	21	338	4.68	21
07-SEP-1999	99US-0152363			338	4.68	21	338	4.68	21
10-SEP-1999	99US-0153070			338	4.68	21	338	4.68	21
13-SEP-1999	99US-0153758			338	4.68	21	338	4.68	

PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0132856.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138647.  
PR 14-JUN-1999; 99US-0138119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144634.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147300.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147933.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159337.  
PR 14-OCT-1999; 99US-0159368.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 44.4%; Score 779.5; DB 21; Length 332;  
Best Local Similarity 49.7%; Pred. No. 1e-71;

Matches 160; Conservative 44; Mismatches 115; Indels 3; Gaps 3;

OY 13 ICMIFLCPLNLSASQRCNNNDKQALLOITKLTALKNPTINDSVSDDDCGMDLVEC-DETS 71  
Db 11 IIfIfIfIttctscldcnqknkllkikkslnpyhaswdpgtdcswwclegcdatv 70  
OY 72 NRIISLIQDDEALTGQIPQVGDLPYLQALMFRKLPMLFGKIPFEISALDKLSLRSS 131  
Db 71 nhvrtaltifsgqisgqpaevgdplyetlvtcklnltqitqtlaklnlmrlstr 130  
OY 132 TSLGAPVLPFPOLTKLTKLCLDSFNKLGIVIPOLSTIPNKAHLERNELETGEIPDIFG 191  
Db 131 lnlgtlpdltfsgqknlelelslndlsqslstpkllalelsrnlktslgsipesfg 190  
OY 192 NFAGS-PDIYLSHNLGTFVPTFARADPIRLDPSGNRLGEGDIFLFGPKRLMLDPSG 250  
Db 191 sfpgtydrltrshnqslgspksgnldfnridsrnkligdaamlsgnktwtslstr 250  
OY 251 NVLSFNSRVOEFPSPSLTYLDLHNQISLSSESLAKLDOTFNVSNNLCGKIPITGNTL 310  
Db 251 nmfgfdiskv-dlptkltgldlnhngltgnipvgwteaplgfntvsnkicoghptgskl 309  
OY 311 QRFDRYAVLHNHSCGAPLPEC 332  
Db 310 qtfdsysyfmhkcicgappleic 331

RESULT 12

AAG05228  
ID AAG05228 standard; Protein; 339 AA.

AC AAG05228;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 1555.

KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

OS Arabidopsis thaliana.

PN EPI033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.  
PF 25-FEB-1999; 99US-0121825.  
XX 25-FEB-1999; 99US-0123180.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135253.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0136629.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.

PR 24-SEP-1999; 99US-0155659.  
PR 26-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 16-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 44.4%; Score 779.5; DB 21; Length 339;  
Best Local Similarity 49.7%; Pred. No. 1.1e-71;  
Matches 160; Conservative 44; Mismatches 115; Indels 3; Gaps 3;

QY 13 ICMIFLCLPNLSASQRCNNNDKALLQIKALKKPTITDSVSDCCGMDVDEC-DETS 71  
DB 18 lllflftlctclskdlnqndklllklkkslnmpnylaswdpqtdccsyclecgdatv 77  
QY 72 NRIISLIQDDEALTGQIPQVGDPLVLOALWPKRLPNLFGKIPKEETISALKDKLSRLSS 131  
DB 78 nhrvtaltltsqslsgslpaeavgdlpylectivtklslnltqtlqtlaklknltmrlslw 137  
QY 132 TSLSGPYLPFPOLTKLTLCDLSFNKLLGVIPOLSTLPNKLALHERNELTGEIPDIFG 191  
DB 138 tnlgtlpdftlsqknlkelefelsndslgspslslclpkllalelslmnltslgspsf 197  
QY 192 NEFAGS-PDITLSHNOITGVPKTFARADPIRLDFSGNRLEGDISLFLFGPKKRLMDLDFSG 250  
DB 198 sfpgtyvpdlrlnhmqslsgslpkslgnldfnrlidstrnlklsqdasmlfgsnkltwslrls 257  
QY 251 NVLSFNFSRVQEPSPSLTYDLNHNQISGLSSEFLAKLDQTFNVSDNNCGKITPGGNL 310  
DB 258 nmfgrldlskv-dlpkcltglldlnmgltcgnlpvwqteapljfflnvsnkicgnlpvqgkl 316  
QY 311 QRPDRAYLHNSCLGAPLPEC 332  
DB 317 qtfdsysyfnkclcgaplelc 338

RESULT 13  
AAG23196  
ID AAG23196 standard; Protein; 325 AA.  
XX  
AC AAG23196;  
XX



DT 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 26412.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hydrolisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX Arabidopsis thaliana.  
OS  
XX EPI033405-A2.  
PN  
XX 06-SEP-2000.  
PD  
XX 25-FEB-2000; 2000EP-0301439.  
PR  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135124.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 27-MAY-1999; 99US-0136021.  
PR 28-MAY-1999; 99US-0136392.  
PR 01-JUN-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 26-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140981.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145951.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.

```
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151080.
PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151338.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 25-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161992.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.
PR 29-OCT-1999: 99US-0162142.
```

Query Match 40.9%; Score 718.5; DB 21; Length 325;

Best Local Similarity 47.2%; Pred. No. 1.9e-65;

Matches 151; Conservative 54; Mismatches 106; Indels 9; Gaps 6;

```
15 MFELCPNLSASQRCNNKQALLOITKLTALKNPTITDSWVSDDDCC-GNDVLECDTSNR 73
13 IILFILP-----SSYSCENCKNALIQLKALGNPPLISWNPRTCCGTCVTECC--tnt 67
74 IISLIQDDEALTGQIPPOVGDLPYLQALMFRKLPNLEFGKPEEISALDLSRLSTS 133
68 vevglsvtsgv-vsgqsyatgldvdrldsfyphltgpnrltclknlnltlykhs 126
134 LSGPVPLEFPQLTKLCLDLSFNKLGIVPPOLSTIPNLKALHLENELTGEIPDIFGNF 193
127 lsgpvpdyiselkrlfildistnftgcpipglsqmpkleaqlindnkltsipnsfgsf 186
194 ACS-PDIYSHNOULTGFVFKTFARADPRLDFSGNRLDEGDISFLFGPKRRLMLDFSGNV 252
187 vgnvpolyismnklsgkipesiskydfnaavdlsgngfegdaafmfigrnktctvrldstrm 246
253 LSFNFSGVGEFPPSLTYLDLNNHQISGSLSSLELAKLDLOTFNVSDDNLCGKIPTGCGNLR 312
```

```
DB 247 fntdlkvv-kfarsivldisqnhlygkipalclklnhfnvsdnhlgckipsqgllqt 305
QY 313 FDRATVLIHNSCLCGAPLPEC 332
DB 306 fepsafehniclcpplkac 325
```

#### RESULT 14

```
AAR37817
ID AAR37817 standard; Protein; 342 AA.
AC AAR37817;
DT 01-OCT-1993 (first entry)
DE Sequence of fungine endopolysialacturonase inhibitor (PGIP).
KW Fungine endopolysialacturonase inhibitor (PGIP); PCR; primer;
KM pesticide.
XX Phaseolus vulgaris, strain Saxa.
OS
PN WO9311241-A.
PD 10-JUN-1993.
XX
PF 04-DEC-1992: 92WO-IT00158.
XX
PR 06-DEC-1991: 91IT-OR00915.
XX
PA (CNDR ) CONSIGLIO NAZ DELLE RICERCHE.
PA (UYGE-) UNIV GEORGIA RES FOUND INC.
PI Albersheim P, Bergmann C, Cervone F, Darvill A;
PI De Lorenzo G, Salvi G;
XX
DR WPI: 1993-197059/24.
DR N-PSDB: AA042595.
XX
PT Nucleotide sequence contained in recombinant vector - encodes
PT endo-polygalacturonase inhibitor (PGIP), useful for conferring
PT e.g. fungal resistance to plants
XX
PS Claim 16: Page 18-19; 36pp; English.
XX
CC The PGIP is capable of inhibiting activity of the fungine
CC endo-alpha-1,4-D-poly-galacturonase enzyme (PG). PGIP is useful
CC in plant transformation to confer fungal resistance or resistance
CC to micrororganisms pathogenicity.
XX
SQ Sequence 342 AA:
```

Query Match 35.4%; Score 622; DB 14; Length 342;

Best Local Similarity 41.0%; Pred. No. 1.8e-55;

Matches 141; Conservative 56; Mismatches 127; Indels 20; Gaps 7;

```
QY 2 NIESFPCILICIMFLCPNLSASQRCNNNDKQALLOITKLTALKNPTITDSWVSDDDCG 61
DB 5 nlpvtmssslilvlalvslreltselcnpgdqalqlkkdignptltsswiprtccn 64
QY 62 --WDVLECCETSN--RIISLIQDDEALTGQ-----IPPOVGDLPLYQALMFRKLPNV 110
DB 65 rtwlgvldctdctqyrvnvl-----disghmlpkypipssianlpynlnflyigginl 118
QY 111 FCKIPEEISALDLSRLSTSLSGVPPLFPQLTKLCLDLSFNKLGIVPPOLSTIP 170
DB 119 vgpilpataklqlylylthcnvsgalpdflsqiklytlldsfynalsqtlppsislsp 178
QY 171 NUKALHLENELTGEIPDIFGNFAG-SPDIYSHNOULTGFVFKTFARADPRLDFSGNRL 229
DB 179 nlgaltfdgnrtisgaldpdygsfskfltamltisrnlcgkippftlanlnlatvdlstrml 238
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2002, 20:14:09 ; Search time 22.58 Seconds

(without alignments)  
330,872 Million cell updates/sec

Title: US-09-308-140-7

Perfect score: 1756

Sequence: 1 MNISSEFCPLICMIFLCL.....FDRATVHLNSCLCAPLPEC 332

Scoring table:

BLOSUM62

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: Issued Patents AA:\*  
2: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep:\*  
7: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	881.5	50.2	330	1	US-08-238-163-2
2	844	48.1	327	1	US-08-238-163-4
3	632	35.4	342	1	US-08-244-646-15
4	632	35.4	342	1	US-08-592-9368-21
5	622	35.4	342	2	US-09-111-573-21
6	464.5	26.5	227	1	US-08-244-646-17
7	322.5	18.4	968	4	US-09-180-439-4
8	320.5	18.3	968	4	US-09-180-439-3
9	319.5	18.2	1016	4	US-09-180-439-8
10	315.5	18.0	1196	4	US-08-881-706-2
11	315	17.9	1112	4	US-09-353-585-2
12	315	17.9	1112	4	US-09-353-585-3
13	311.5	17.7	999	4	US-08-473-553A-5
14	300.5	17.1	799	4	US-09-180-439-6
15	300.5	17.1	1012	2	US-08-473-553A-4
16	300.5	17.1	1025	2	US-08-567-375-4
17	300.5	17.1	1025	2	US-08-587-680A-4
18	283	16.1	863	4	US-08-666-271-2
19	278.5	15.9	806	4	US-08-945-983-2
20	273	15.5	980	2	US-08-473-553A-5
21	273	15.5	985	2	US-08-473-553A-2
22	264	15.0	523	2	US-08-473-553A-3
23	226.5	12.9	1023	2	US-08-475-891A-2
24	226.5	12.9	1023	2	US-08-567-375-2
25	226.5	12.9	1023	2	US-08-587-680A-2
26	188	10.7	268	4	US-09-353-585-6
27	172	9.8	301	4	US-09-353-585-5

28	172	9.8	544	2	US-08-587-680A-25	Sequence 25, Appl
29	155	8.8	277	2	US-08-567-375-16	Sequence 16, Appl
30	154.5	8.8	375	1	US-08-303-238-2	Sequence 2, Appl
31	154.5	8.8	375	4	US-08-458-834-2	Sequence 2, Appl
32	149.5	8.5	605	4	US-09-063-950-5	Sequence 5, Appl
33	147	8.4	282	1	US-08-442-063A-45	Sequence 45, Appl
34	147	8.4	307	1	US-08-442-063A-48	Sequence 48, Appl
35	147	8.4	333	1	US-08-442-063A-27	Sequence 27, Appl
36	147	8.4	342	1	US-08-272-919-2	Sequence 2, Appl
37	147	8.4	342	1	US-08-619-916-2	Sequence 2, Appl
38	147	8.4	342	5	PCT-US95-08542-2	Sequence 2, Appl
39	147	8.4	359	1	US-08-303-238-4	Sequence 4, Appl
40	147	8.4	359	4	US-08-458-834-4	Sequence 4, Appl
41	146.5	8.3	605	1	US-08-190-802A-49	Sequence 49, Appl
42	146.5	8.3	605	4	US-08-477-346-49	Sequence 49, Appl
43	146	8.3	376	1	US-08-303-238-1	Sequence 1, Appl
44	146	8.3	376	4	US-08-458-834-1	Sequence 1, Appl
45	144.5	8.2	603	1	US-08-190-802A-50	Sequence 50, Appl

#### ALIGNMENTS

RESULT 1  
US-08-238-163-2  
Sequence 2, Application US/08238163  
Patent No. 5569830  
GENERAL INFORMATION:  
APPLICANT: BENNETT, Alan  
APPLICANT: LABAVITCH, John M.  
APPLICANT: FOWELL, Ann  
APPLICANT: STOTZ, Henrik  
TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL  
TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/238,163  
FILING DATE: 03-MAY-1994  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 2307E-540  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 330 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-238-163-2

Query Match 50.2%; Score 881.5; DB 1; Length 330;  
Best Local Similarity 54.7%; Pred. No. 7.1e-88;  
Matches 181; Conservative 41; Mismatches 106; Indels 3; Gaps 3;  
QY 3 ISSSEFCPLICMIFLCLPNLSASORCNNDKQALQIKTALKNPTITDSDWSDDCGGM 62



Query Match 35.4%; Score 622; DB 1; Length 342;

Best Local Similarity 41.0%; Pred. No. 1.5e-59; Indels 20; Gaps 7;

Matches 141; Conservative 56; Mismatches 127;

2 NISSFCPLICMIFLCIPNLASORCANNNDKQALLQITKLTALKNPTTIDSWVSDCCG 61  
 5 NIPVMSSSLITLVLSRTALSELCPNODKQALLQITKLTALKNPTTIDSWVSDCCG 64  
 62 --WDVECOETSN--RIISLIQDDEALTGQ-----IPQVGDLPYQALMFRKLP 110  
 65 RTWGLVLCPTDQTYRVNML-----DLGHNLPKPYIPSSLANPYLNFLYIGGINNL 118  
 111 FCKIPEISALDKLSRLSTSLSGPVLPFPQTLKLCDSLFPKLGIVIPOLSTIP 170  
 119 VGPPIPAIAKLQHLHYLTHTNVSGAIPDLSQIKTLVTLDFSYNALSGTLPPSISLP 178  
 171 NIKALHLERNELTGEIPDIFGNFAG-SPDIYLSHNOITGFEVPTKTRADPIRLDFSGNRL 229  
 179 NIGGITFDGNRISGAIPDSYGSFKLTAMTISRNLTKIPPTFANLALAVDLSRNML 238  
 230 EGDISFLFQPKRLEMLDFSGNVLSFNSRVOEFPPLTYLDLNNHSGLSSELA 289  
 239 EGDASVLFSGDKTKIKHLAKNSLAFDLGKV-GLSKNLGLDLRNNRIYGTLPQGLTOLK 297  
 290 -LQTFVSDNNLCGRIPITGNNLQRPRTAYLHNSCLCGAPLPEC 332  
 298 FLOSLNVSFNNLCGEIPQGNLKRFDVSSYANNKCLCGSPPLPSC 341

## RESULT 4

US-08-592-936B-21

Sequence 21, Application US/08592936B

Patent No. 5783393

GENERAL INFORMATION:

APPLICANT: Kelllogg, Jill A.

TITLE OF INVENTION: PLANT TISSUE / STAGE SPECIFIC PROMOTERS FOR

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESS: Dehlinger & Associates

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/592,936B

FILING DATE: 29-JAN-1996

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Evans, Susan T.

REGISTRATION NUMBER: 38,443

REFERENCE/DOCKET NUMBER: 4257-0012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 324-0880

TELEFAX: (650) 324-0960

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 342 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHEICAL: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: predicted amino acid coding sequence

INDIVIDUAL ISOLATE: OF SEQ ID NO:20

US-08-592-936B-21

Query Match 35.4%; Score 622; DB 1; Length 342;

Best Local Similarity 41.0%; Pred. No. 1.5e-59; Indels 20; Gaps 7;

Matches 141; Conservative 56; Mismatches 127;

2 NISSFCPLICMIFLCIPNLASORCANNNDKQALLQITKLTALKNPTTIDSWVSDCCG 61  
 5 NIPVMSSSLITLVLSRTALSELCPNODKQALLQITKLTALKNPTTIDSWVSDCCG 64  
 62 --WDVECOETSN--RIISLIQDDEALTGQ-----IPQVGDLPYQALMFRKLP 110  
 65 RTWGLVLCPTDQTYRVNML-----DLGHNLPKPYIPSSLANPYLNFLYIGGINNL 118  
 111 FCKIPEISALDKLSRLSTSLSGPVLPFPQTLKLCDSLFPKLGIVIPOLSTIP 170  
 119 VGPPIPAIAKLQHLHYLTHTNVSGAIPDLSQIKTLVTLDFSYNALSGTLPPSISLP 178  
 171 NIKALHLERNELTGEIPDIFGNFAG-SPDIYLSHNOITGFEVPTKTRADPIRLDFSGNRL 229  
 179 NIGGITFDGNRISGAIPDSYGSFKLTAMTISRNLTKIPPTFANLALAVDLSRNML 238  
 230 EGDISFLFQPKRLEMLDFSGNVLSFNSRVOEFPPLTYLDLNNHSGLSSELA 289  
 239 EGDASVLFSGDKTKIKHLAKNSLAFDLGKV-GLSKNLGLDLRNNRIYGTLPQGLTOLK 297  
 290 -LQTFVSDNNLCGRIPITGNNLQRPRTAYLHNSCLCGAPLPEC 332  
 298 FLOSLNVSFNNLCGEIPQGNLKRFDVSSYANNKCLCGSPPLPSC 341

## RESULT 5

US-09-111-573-21

Sequence 21, Application US/09111573

Patent No. 5929302

GENERAL INFORMATION:

APPLICANT: Kelllogg, Jill A.

TITLE OF INVENTION: PLANT TISSUE / STAGE SPECIFIC PROMOTERS FOR

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESS: Dehlinger & Associates

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/111,573

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/592,936

FILING DATE: 29-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Evans, Susan T.

REGISTRATION NUMBER: 38,443

REFERENCE/DOCKET NUMBER: 4257-0012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 324-0880

TELEFAX: (650) 324-0960

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 342 amino acids

TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEetical: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: predicted amino acid coding sequence  
INDIVIDUAL ISOLATE: of SEQ ID NO:20  
US-09-111-573-21

Query Match 35.4%; Score 622; DB 2; Length 342;  
Best Local Similarity 41.0%; Pred. No. 1.5e-59;  
Matches 141; Conservative 56; Mismatches 127; Indels 20; Gaps 7;

Db 2 NIESSFCPIICMIFLCLPNLSASQRCNNNDKXALLQIKALKNPITITDSVWSDDCG 61  
5 NIPVWSSSLITVLVSRLTALSELCPQDKQALLQIKDLGPTLSSLPPTTDCN 64  
62 --WDLVECDSTN--RIYSLIIQDEALTGQ-----IPQVGDLPYLQALMFRKLPNL 110  
65 RTWGLVLCDTQGTQYRVNML-----DLSGHNLPKPYPIPSLALPFLNLYIGGINNL 118  
111 FGKIPPEISALKDLKSLRSLSTSLSGVPVLPFPQTLTKLCLDLSEFKLLGYIPQSLTP 170  
119 VGPPIPAIAKLTQLHYLYITHTNWSGAIPEFLSQIKTLVTLDFSYNALSGTLPPSISLP 178  
171 NLKALHLERNELTGEIPDLFGNFAG--SPDIYLSHQULTGFVFKTFARADPIRLDPSGNRL 229  
179 NLGGTFEGNRIISGAIIPDSYGSFSLFTAMTISRRLTKIPPTFANLNLAFLVDLSRNL 238  
230 EGDISFLGPKKRLKLEMLDFSGVNLSEFNSRVOEFPPLSTLYLDLNHQISLSSELA 289  
239 EGDASVFLGSDKNTKIKILAKNSLAFDLGV--GLSKNLNGDLRNNRITGTLPQGLTQ 297  
290 -LQTFNVSNNLCKIPTGNNLQRPDRTAYLHNSCLGAPLPEC 332  
298 FLOSINVSFNNLCEIPQGNLKRFDVSSYANNKCLCGSPLPSC 341

RESULT 6  
US-08-244-646-17  
Sequence 17, Application US/08244646  
Patent No. 5744692  
GENERAL INFORMATION:  
APPLICANT: Cervone, Felice  
APPLICANT: De Lorenzo, Giulia  
APPLICANT: Salvi, Giovanni  
APPLICANT: Albersheim, Peter  
APPLICANT: Darvill, Alan  
APPLICANT: Bergmann, Carl  
TITLE OF INVENTION: Nucleotide Sequences Coding An  
TITLE OF INVENTION: Endopolygalacturonase Inhibitor  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sally A. Sullivan  
STREET: 5370 Manhattan Circle Suite 201  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/244,646  
FILING DATE: 06-JUN-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IT RM 91A 000915  
FILING DATE: 06-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/IT/00158

FILING DATE: 04-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Sullivan, Sally A.  
REGISTRATION NUMBER: 32,064  
REFERENCE/DOCKET NUMBER: 19-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303)499-8080  
TELEFAX: (303)499-8089  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 227 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-244-646-17

Query Match 26.5%; Score 464.5; DB 1; Length 227;  
Best Local Similarity 43.4%; Pred. No. 1.2e-42;  
Matches 98; Conservative 41; Mismatches 84; Indels 3; Gaps 3;

Db 109 NLFKIPPEISALKDLKSLRSLSTSLSGVPVLPFPQTLTKLCLDLSEFKLLGYIPQSLTP 168  
2 NLVGPPIPAIAKLTQLHYLYITHTNWSGAIPEFLSQIKTLVTLDFSYNALSGTLPPSISLP 61  
169 LPNKALHLERNELTGEIPDLFGNFAG--SPDIYLSHQULTGFVFKTFARADPIRLDPSGN 227  
62 LPNVIGTFIDNRIISGAIIPDSYGSFSLFTSMTISRRLTKIPPTFANLNLAFLVDLSRNL 121  
228 RLEDISFLGPKKRLKLEMLDFSGVNLSEFNSRVOEFPPLSTLYLDLNHQISLSSELA 287  
122 MLOGDASVFLGSDKNTKIKILAKNSLAFDLGV--GLSKNLNGDLRNNRITGTLPQGLTQ 180  
288 LD-LQTFNVSNNLCKIPTGNNLQRPDRTAYLHNSCLGAPLPEC 332  
181 LKPLHSLVSNLCEIPQGNLQTFHVSAYANNKCLCGSPLPAC 226

RESULT 7  
US-09-180-439-4  
Sequence 4, Application US/09180439  
Patent No. 6225532  
GENERAL INFORMATION:  
APPLICANT: Dixon, Mark S  
APPLICANT: Hatzixanthis, Kostas  
APPLICANT: Jones, David A  
APPLICANT: Jones, Jonathan DG  
TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof  
FILE REFERENCE: 620 - 53  
CURRENT APPLICATION NUMBER: US/09/180,439  
CURRENT FILING DATE: 1998-12-06  
EARLIER APPLICATION NUMBER: PCT/GB97/01249  
EARLIER FILING DATE: 1997-05-08  
EARLIER APPLICATION NUMBER: GB 9609681.3  
EARLIER FILING DATE: 1996-05-09  
EARLIER APPLICATION NUMBER: GB 9619924.5  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 968  
TYPE: PRT  
ORGANISM: Lycopersicon esculentum  
US-09-180-439-4

Query Match 18.4%; Score 322.5; DB 4; Length 968;  
Best Local Similarity 29.3%; Pred. No. 3.8e-26;  
Matches 111; Conservative 45; Mismatches 142; Indels 81; Gaps 12;  
Db 30 NNNDKQALLQIKALKNP--TITDSVWSDDCG--GMDLVE----- 67  
27 STEATALLWKATFFKNQNSFLASWTTSSNACKDYGVCVCLNGRVNTLNTNASVIGTL 86

QY 68 -----DETSNRI-----ISLIID--DEALTOIIPQVDDPYLQ 100  
DB 87 YAFPFSSLPFLENLDLNNISGTTIPEIGNLTNLVYLDLNTNOISGTTIPOQISLAKLQ 146  
QY 101 ALMFRKLPN-LFGKIPPEISALKDKLSRLSSTLSGVPVLPFPOLTKLTCLDLSPFKLL 159  
DB 147 II--RIFNHNLNGFIPEEIGYLRSLTKSLGINFSLGSIIPASIGNMTNLSFLYENOLS 204  
QY 160 GVIPPOLSTLPNKLALHLELNEGTGEPIDIFGNFAGSPDIYLSHNOULTGVPKTFARADP 219  
DB 205 GFPEEIGYLRSLTKSLDINFSLGSIIPASIGNMTNLSFLYENOLS 264  
QY 220 I-RDPSGNRLEGDISFLPGPKRLEMDPSGNVLSFNFSRVOEFPSTLYDLNHNQIS 278  
DB 265 LTKLSLGINFLSGSIIPASIGNMTNLSRLDLYNNKLSGSIPEEIGYLRSLTYLDGENALN 324  
QY 279 GSLSSELAKL-DLQTFNVSDDNLCGRIP-----TGGNLOR 312  
DB 325 GSIPSSLGNNLNSRLDLYNNKLSGSIPEEIGYLRSLTYLDGENALNGSIIPASIGNLNN 384  
QY 313 FDRATYLNHNSCLGAPLPE 331  
DB 385 L-FMLYLYNNOLSGS-IPE 401  
RESULT 8  
US-09-180-439-3  
; Sequence 3, Application US/09180439  
; Patent No. 6225532  
; GENERAL INFORMATION:  
; APPLICANT: Dixon, Mark S  
; APPLICANT: Hatzixanthis, Kostas  
; APPLICANT: Jones, David A  
; APPLICANT: Jones, Jonathan DG  
; TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof  
; FILE REFERENCE: 620 - 53  
; CURRENT APPLICATION NUMBER: US/09/180,439  
; EARLIER FILING DATE: 1998-12-06  
; EARLIER APPLICATION NUMBER: PCT/GB97/01249  
; EARLIER FILING DATE: 1997-05-08  
; EARLIER APPLICATION NUMBER: GB 9609681.3  
; EARLIER FILING DATE: 1996-05-09  
; EARLIER APPLICATION NUMBER: GB 9619924.5  
; EARLIER FILING DATE: 1996-09-24  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 968  
; TYPE: PRT  
; ORGANISM: Lycopersicon esculentum  
US-09-180-439-3  
Query Match 18.3%; Score 320.5; DB 4; Length 968;  
Best Local Similarity 29.3%; Pred. No. 6.2e-26;  
Matches 111; Conservative 44; Mismatches 143; Indels 81; Gaps 12;  
QY 30 NNNDKQALLQIKTALKNP--TITDSWVSDDCC-GWDLVE----- 67  
DB 27 STEBATALLTKKATFKKQNNNSFLASWTSSNACKDWGVCLNGRVNTLNTNASVIGTL 86  
QY 68 -----DETSNRI-----ISLIID--DEALTOIIPQVDDPYLQ 100  
DB 87 YAFPFSSLPFLENLDLNNISGTTIPEIGNLTNLVYLDLNTNOISGTTIPOQISLAKLQ 146  
QY 101 ALMFRKLPN-LFGKIPPEISALKDKLSRLSSTLSGVPVLPFPOLTKLTCLDLSPFKLL 159  
DB 147 II--RIFNHNLNGFIPEEIGYLRSLTKSLGINFSLGSIIPASIGNMTNLSFLYENOLS 204  
QY 160 GVIPPOLSTLPNKLALHLELNEGTGEPIDIFGNFAGSPDIYLSHNOULTGVPKTFARADP 219  
DB 205 GFPEEIGYLRSLTKSLDINFSLGSIIPASIGNMTNLSFLYENOLS 264

QY 220 I-RDPSGNRLEGDISFLPGPKRLEMDPSGNVLSFNFSRVOEFPSTLYDLNHNQIS 278  
DB 265 LTKLSLGINFLSGSIIPASIGNMTNLSRLDLYNNKLSGSIPEEIGYLRSLTYLDGENALN 324  
QY 279 GSLSSELAKL-DLQTFNVSDDNLCGRIP-----TGGNLOR 312  
DB 325 GSIPSSLGNNLNSRLDLYNNKLSGSIPEEIGYLRSLTYLDGENALNGSIIPASIGNLNN 384  
QY 313 FDRATYLNHNSCLGAPLPE 331  
DB 385 L-FMLYLYNNOLSGS-IPE 401  
RESULT 9  
US-09-180-439-8  
; Sequence 8, Application US/09180439  
; Patent No. 6225532  
; GENERAL INFORMATION:  
; APPLICANT: Dixon, Mark S  
; APPLICANT: Hatzixanthis, Kostas  
; APPLICANT: Jones, David A  
; APPLICANT: Jones, Jonathan DG  
; TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof  
; FILE REFERENCE: 620 - 53  
; CURRENT APPLICATION NUMBER: US/09/180,439  
; EARLIER FILING DATE: 1998-12-06  
; EARLIER APPLICATION NUMBER: PCT/GB97/01249  
; EARLIER FILING DATE: 1997-05-08  
; EARLIER APPLICATION NUMBER: GB 9609681.3  
; EARLIER FILING DATE: 1996-05-09  
; EARLIER APPLICATION NUMBER: GB 9619924.5  
; EARLIER FILING DATE: 1996-09-24  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 1016  
; TYPE: PRT  
; ORGANISM: Lycopersicon esculentum  
US-09-180-439-8  
Query Match 18.2%; Score 319.5; DB 4; Length 1016;  
Best Local Similarity 29.0%; Pred. No. 8.6e-26;  
Matches 110; Conservative 45; Mismatches 143; Indels 81; Gaps 12;  
QY 30 NNNDKQALLQIKTALKNP--TITDSWVSDDCC-GWDLVE----- 67  
DB 27 STEBATALLTKKATFKKQNNNSFLASWTSSNACKDWGVCLNGRVNTLNTNASVIGTL 86  
QY 68 -----DETSNRI-----ISLIID--DEALTOIIPQVDDPYLQ 100  
DB 87 YAFPFSSLPFLENLDLNNISGTTIPEIGNLTNLVYLDLNTNOISGTTIPOQISLAKLQ 146  
QY 101 ALMFRKLPN-LFGKIPPEISALKDKLSRLSSTLSGVPVLPFPOLTKLTCLDLSPFKLL 159  
DB 147 II--RIFNHNLNGFIPEEIGYLRSLTKSLGINFSLGSIIPASIGNMTNLSFLYENOLS 204  
QY 160 GVIPPOLSTLPNKLALHLELNEGTGEPIDIFGNFAGSPDIYLSHNOULTGVPKTFARADP 219  
DB 205 GFPEEIGYLRSLTKSLDINFSLGSIIPASIGNMTNLSFLYENOLS 264  
QY 220 I-RDPSGNRLEGDISFLPGPKRLEMDPSGNVLSFNFSRVOEFPSTLYDLNHNQIS 278  
DB 265 LTKLSLGINFLSGSIIPASIGNMTNLSRLDLYNNKLSGSIPEEIGYLRSLTYLDGENALN 324  
QY 279 GSLSSELAKL-DLQTFNVSDDNLCGRIP-----TGGNLOR 312  
DB 325 GSIPSSLGNNLNSRLDLYNNKLSGSIPEEIGYLRSLTYLDGENALNGSIIPASIGNLNN 384  
QY 313 FDRATYLNHNSCLGAPLPE 331  
DB 385 LSRDL-LYNNKLSGS-IPE 401



RESULT 10  
US-08-881-706-2  
Sequence 2, Application US/08881706  
Patent No. 6245969  
GENERAL INFORMATION:  
APPLICANT: Chong, Joane  
APPLICANT: Li, Jianming  
TITLE OF INVENTION: Receptor Kinase BIN1  
FILE REFERENCE: 07251/022001  
CURRENT APPLICATION NUMBER: US/08/881,706  
CURRENT FILING DATE: 1997-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1196  
TYPE: PRT  
ORGANISM: Arabidopsis sp.  
US-08-881-706-2

Query Match 18.0%; Score 315.5; DB 4; Length 1196;  
Best Local Similarity 28.3%; Pred. No. 3.1e-25;  
Matches 112; Conservative 48; Mismatches 143; Indels 93; Gaps 13;  
QY 20 LPTNSAS-----QRCNNNDKQALLQIKALKNP--TITDSWVSDDDCCGW--DLVECDET 70  
DB 385 LTNLSASLRLTLDDSSNFGSPIL--PMLCONPKNTLOELYLQNNFTGKIPPTLSNCSE- 441  
QY 71 SNRIISLIIDDEALTGQIPPOVGDLPYLO--ALMFRKLPLNFGKIPPEEISALKDLKSLR 128  
DB 442 ---LVSLHLSEFN-VLSTGTISSLSGLSKLDKMLMLE---GEIQELMYVYKLTETLI 494  
QY 129 LSSTSGSPVPLFPQTLTKTCLDLSFNKLLGVIPQLSTPLNKAHLERNELTGEIPD 188  
DB 495 LDFENDLTGEIPSGISLNCNTLNMIWISLNNRLTGEIPKWIGRLNLAIKLSNNSGSIIPD 554  
QY 189 IFGNFASPDILYSHNOLTFGVPT--FARADPTKLPD-----SGN 227  
DB 555 ELGGCRSLMLDLNTLNTFNCTIPAMFKQSGKIAANFIAGKRYVYIKNDMKKECHGAGN 614  
QY 228 RLE-----DISPLFGPKRRLKLELDPSGNVLSFNSRVQ 261  
DB 615 LLEFGIRSDOLANLSTRNRCNITSRYVGHSTPTFDNNSMFLDMSTYMLSGYIKETI 674  
QY 262 EFPSPLYLDLNNHNOISGLSSELA-----KLD-----LOTFNVS 296  
DB 675 GSPMYLFLNLNGHNDISGSIPEVGDRLGLNILDLSNKLGDRIPOAMSA/LTMLTEIDLS 734  
QY 297 DNNLCGKIPTGNGNQRDRATAYLHNSCLGAPLPEC 332  
DB 735 NNNLSGPIPEMGOFETFPAPKFLNNPGLCGYPLPRC 770

RESULT 11  
US-09-353-585-2  
Sequence 2, Application US/09353585  
Patent No. 6287865  
GENERAL INFORMATION:  
APPLICANT: Dixon, Mark S  
APPLICANT: Jones, David A  
APPLICANT: Jones, Jonathan DG  
TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye PC  
STREET: 8th Floor, 1100 No. 6287865th Glebe Road  
CITY: Arlington  
STATE: Virginia  
COUNTRY: United States of America  
ZIP: 22201-4714

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/353,585  
FILING DATE: 15-Jul-1999  
CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12O 1/68  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/930,277  
FILING DATE: 27-OCT-1997  
APPLICATION NUMBER: PCT/GB96/00785  
FILING DATE: 01-APR-1996  
APPLICATION NUMBER: GB 950658.5  
FILING DATE: 31-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ms. Mary J. Wilson  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 620-69  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1112 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Tomato  
STRAIN: Cf2  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-353-585-2

Query Match 17.9%; Score 315; DB 4; Length 1112;  
Best Local Similarity 30.8%; Pred. No. 3.1e-25;  
Matches 108; Conservative 41; Mismatches 146; Indels 56; Gaps 11;  
QY 30 NNNDKQALLQIKALKNP--TITDSWV--SDDDCCGMDLYVC----- 67  
DB 27 STEEATALLKMKATFFKQNNSEFLASWIPSSNACKDYGVCFGNRYNTLNTASVIGTL 86  
QY 68 -----DEFSNRI-----ISLIID--DEALTGQIPPOVGDLPYLO 100  
DB 87 YAFPPSSLPSENLDSLKNNTYGTIPPEIGNLTNLVYLDLNNNOISGTTIPQIGLAKID 146  
QY 101 ALMFRKLPN-LFGKIPPEISALKDLKSLRSLSTSGSPVPLFPQTLTKTCLDLSFNKLL 159  
DB 147 II--RIFHNLQNGFIPEKIGYLRSLTKLSIGINFLSGSIPASVGNLNNLSFLYLYNNQLS 204  
QY 160 GVIPQLSTPLNKAHLERNELTGEIPDIFGNFASPDILYSHNOLTFGVPTFARAPD 219  
DB 205 GSIPPEISTYLRSLTELDLSDNALNGSIPASLGMMNLSFLFLGNQLSGISPEICYLRS 264  
QY 220 IR-LDFSGNRLEGDISFLFGPKRRLKLELDPSGNVLSFNSRVQEFPPSLTYLDLNNHNOIS 278  
DB 265 LTYLDLSENALNGSIPASLGNNLNSFLFLYGNQLSGISPEICYLRS/LNVLGSLSEMAIN 324  
QY 279 GSISSFLAKV-LDQTFNVSDNNLCGKIPTG--GNLQRFDRATAYLHNSCLGCA 327  
DB 325 GSIPASLGNLKLNSRLNLVNNQLSGISIPASLGNLNLNS--MLYLYNNQLSGS 374

RESULT 12  
US-09-353-585-3  
Sequence 3, Application US/09353585  
Patent No. 6287865  
GENERAL INFORMATION:

```

APPLICANT: Dixon, Mark S
            Jones, David A
            Jones, Jonathan DG
TITLE OF INVENTION: Plant pathogen resistance genes and uses
thereof
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: Nixon & Vanderhye PC
STREET: 8th Floor, 1100 No. 6287865th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/353,585
FILING DATE: 15-Jul-1999
CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12O
1/68
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/930,277
FILING DATE: 27-OCT-1997
APPLICATION NUMBER: PCT/GB96/00785
FILING DATE: 01-APR-1996
APPLICATION NUMBER: GB 9506658.5
FILING DATE: 31-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ms Mary J Wilson
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 620-69
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Tomato
STRAIN: Cf2
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-353-585-3

```

```

Query Match 17.9%; Score 315; DB 4; Length 1112;
Best Local Similarity 30.8%; Pred. No. 3 1e-25;
Matches 108; Conservative 41; Mismatches 146; Indels 56; Gaps 11;
30 NNNDKALLQITALKNP--TTDSWV-SDDCCGMDLVE----- 67
27 STEEATALLKMKATEFNQNSFLASWIPSSNACKOWYGVCFGRVNTLNTNASYIGTL 86
68 -----DETSRI-----ISLIID--DEALTGOIPPOVGDLPVQ 100
87 YAFPPSSLPSELENDLSKNITGTPPEIGNTLNLVYLDLNNNOISGTPQIGLAKIQ 146
101 ALMFKLPN-LFGKIPPEISALKDKLSRLSTSLGVPPLFPOLTKTCDISFNKL 159
147 IL--RIFHNQNLGFIPEKELGYLSLTKLSLGINFLSGSIPASVGNLNNLSFLYLNQIS 204
160 GYIPPOLSTPLMKALHLERNELTGEIPDIFGNFAGSPDIYLSHNLGTGFVKTPARADP 219
205 GSIPPEISYLSRLTEILSDNALNGSIPASLGNMNNLSFLFYGNQSGSIPPEELCYLRS 264
220 IR-LDFSGRLGSDISLFGPKKRLKLEMDFSGNVLSFNFRAQVEPPSLTYLDLNNQIS 278

```

---

```

DB 265 LTYLDLSENALNGSIPASIGNLNLSFLFYGNQSGSIPPEIGYLSRLNLVGLSEMLN 324
OY 279 GSLSLELKL-LDLPFNVSDDNLCKIPTG--GNLORFDRATVLAHSCICGA 327
DB 325 GSIPASLGNLKLRLNLVNNQSGSIPASLGNLNLNLS-MLYLYNNQISGS 374

```

RESULT 13  
US-08-473-553A-5  
Sequence 5, Application US/08473553A  
Patent No. 5859338  
GENERAL INFORMATION:  
APPLICANT: Meyerowitz, Elliot M.  
APPLICANT: Clark, Steven E.  
APPLICANT: Williams, Robert W.  
TITLE OF INVENTION: Plant Clavatal Nucleic Acids,  
TITLE OF INVENTION: Transformed Plants, and Proteins  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,553A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Silva, Robin M.  
REGISTRATION NUMBER: 38,304  
REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 999 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-473-553A-5

---

```

Query Match 17.7%; Score 311.5; DB 2; Length 999;
Best Local Similarity 26.6%; Pred. No. 6.3e-25;
Matches 115; Conservative 49; Mismatches 125; Indels 143; Gaps 16;
7 FCPILCICIFLCPLNLSQRCNNNDKQALLQITALKNPIT--DSWVSDDC--CGWD 63
DB 3 YCLILLCISLYLPSLSL--NODATILRQAKLGSDPAQSSLSWSMDNDVTPCKKL 57
OY 64 LVECDETS-----RIL 75
DB 58 GVSQATSNVSVSDLSFNLVGPSPSILCHPLSLSLSYNNSINGSLASADFDCHNLI 117
OY 76 SLIIDDEALTGOIPPOVG--DLPYQAL-----WFKLP----- 108
DB 118 SLDLSEN-LVGSIPKSLPFPNLPNLFLEISGNNSDTPSPSGFGRKLSLNLAGNLS 176
OY 109 -----NLF--GKIPPEISALKDKLSRLSTSLGVPPLFPOL 146
DB 177 GTIPASLGNVTLTKELKLAVALNLFSPSQIPQOLGNLTGLVLAAGCNVCPIPPSLRIT 236

```



```

QY 216 -----RADPIR-----LDPSGNRLGDISFLGPKKRL 244
      ||:|
Db 241 PNSIMWLSLRAFSVRENKLGMIPTNAFKTLHLLEVIMGTNRPHKIPASVANASHLT 300
      ||:|
QY 245 MLDPSGN---VLSFNFPSRYOEFPSLTYLDLNNH-----QISGSLSSHAKLD-LQTF 293
      ||:|
Db 301 VIOIYGNLFSGIITSGFGRLR---NLTELYLWRNLFOFREDDMGFISDLTNGSKLQTL 356
      ||:|
QY 294 NVSDNNLGGKIP 305
      ||:|
Db 357 NLGNNLGGVLP 368
      ||:|

```

Search completed: March 13, 2002, 20:56:46  
 Job time: 2557 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2002, 20:16:09 ; Search time 33.71 Seconds

(without alignments)  
750.221 Million cell updates/sec

Title: US-09-308-140-7

Perfect score: 1756  
Sequence: 1 MNISSEFCPLICMIFLCL.....FDRATYLNHNSCLGAPLPEC 332

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	884.5	50.4	330	2 JQ2262	Polygalacturonase
2	849	48.3	327	2 T10263	probable polygalac
3	844	48.1	327	2 S47965	Polygalacturonase
4	819	46.6	329	2 T17033	leucine rich repea
5	632	35.4	342	2 S23764	polygalacturonase-
6	567	32.3	313	2 S60713	polygalacturonase-
7	411.5	23.4	478	2 H86459	hypothetical prote
8	408.5	23.3	371	2 T49908	hypothetical prote
9	401	22.8	480	2 T00971	probable disease r
10	398.5	22.7	512	2 G86459	hypothetical prote
11	378	21.5	976	2 B84659	probable receptor-
12	370	21.1	486	2 B86460	hypothetical prote
13	370	21.1	907	2 A86460	hypothetical prote
14	351	20.0	1232	2 T05322	hypothetical prote
15	348.5	19.8	932	2 T48489	receptor-like prote
16	346	19.7	992	2 T05335	hypothetical prote
17	338	19.2	1064	2 B86465	probable protein k
18	335	19.1	981	2 T50851	receptor protein k
19	330.5	18.8	719	2 T47727	hypothetical prote
20	330.5	18.8	983	2 G84524	probable disease r
21	329.5	18.8	987	2 T50850	receptor protein k
22	327.5	18.7	1002	2 T46033	receptor protein k
23	327.5	18.6	964	2 T49038	hypothetical prote
24	325.5	18.5	905	2 T00475	probable disease r
25	323.5	18.4	1027	2 B85089	receptor protein k
26	322.5	18.4	1013	2 T10659	probable serine/ch
27	322.5	18.3	1120	2 B86479	hypothetical prote
28	322	18.3	988	2 T45717	receptor-kinase 11
29	322	18.3	1143	2 B84431	probable receptor

30	322	18.3	1192	2 T48499	receptor-like prot
31	321.5	18.3	869	2 A71400	probable disease r
32	321.5	18.3	1133	2 E86308	hypothetical prote
33	319.5	18.2	1016	2 T30553	disease resistance
34	317	18.1	702	2 A86383	hypothetical prote
35	316.5	18.0	1123	2 D96756	receptor-like prote
36	315	17.9	967	2 T48210	hypothetical prote
37	315	17.9	1112	2 T10504	disease resistance
38	314.5	17.9	980	2 H84632	probable receptor-
39	313	17.8	1003	2 T05898	hypothetical prote
40	312.5	17.8	766	2 T01817	hypothetical prote
41	312	17.8	1025	2 T45647	receptor protein k
42	311.5	17.7	999	2 S27756	receptor-like prote
43	309.5	17.6	598	2 C96756	receptor-like prote
44	309.5	17.6	960	2 G84652	probable receptor-
45	309	17.6	719	2 T02154	protein kinase hom

## ALIGNMENTS

RESULT 1  
JQ2262  
Polygalacturonase inhibitor precursor - Pyrus communis  
C:Species: Pyrus communis  
C:Date: 04-Sep-1998 #sequence-revision 04-Sep-1998 #text-change 21-Jan-2000  
C:Accession: JQ2262  
R:Stoltz, H.U.; Powell, A.L.T.; Damon, S.E.; Greve, L.C.; Bennett, A.B.; Labavitch, J.  
Plant Physiol. 101, 133-138, 1993  
A:Title: Molecular characterization of a polygalacturonase inhibitor from Pyrus commu  
A:Reference number: JQ2262  
A:Accession: JQ2262  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-330 <STO>  
A:Cross-references: GB:I09264; NID:g169683; PID:AAA33865.1; PID:g169684  
C:Superfamily: polygalacturonase-inhibiting protein; leucine-rich alpha-2-glycoprotei  
F:144-167/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 50.4%; Score 884.5; DB 2; Length 330;

Best Local Similarity 55.0%; Pred. No. 1.1e-64;

Matches 182; Conservative 40; Mismatches 106; Indels 3; Gaps 3;

QY	3	IESSEFCPLICMIFLCLPNLSSORCNNDKQALLQITALKNPITDSWSDDCGW	62
DB	1	MEUKFSTFLSTLTFSSVINPALSDCNPDDKVLQIKKAFEDPVLASMSDSDCCDW	60
QY	63	DVVECDSTNRIITSLTIODEALTQIPQVCDLPIQLALMFRKLPNLFGKIPETISALK	122
DB	61	YCVTCSTTNRIINSLITIFAGQ-VSGQIPALVGDLPYLELLEFFKQPNLNGPIQPAIAKLK	119
QY	123	DKSLRSTLSGVPVLPFPOLTKITCLDSFNKLLGVIPOLSTLPVKAHLHERNEL	182
DB	120	GKSLRSTLSGVPVLPFPOLTKITCLDSFNKLLGVIPOLSTLPVKAHLHERNEL	179
QY	183	TGEIPFIFGNFACS-FDITLSHNQLTGFPVKTFARADPIRLDESGRLEGGDSIFLGPXK	241
DB	180	TGHIPIPSFGPIGNPDVLSHNQLSGNIPISFQMDFTSIDSRNKLGGDSVIFGLXK	239
QY	242	RLMDLSEGNVLSFNSRVOEPPSLTYLDLNNHNSGSLSELAKLDQTFNVSDDNLC	301
DB	240	TTOYIDLNRNLEFNLSKV-EPTSLTSLDINHKNKTYGSIPEVFTOLNFOPLNVSINRLC	298
QY	302	GRIPTGNNLQREDRTAYLHNSCLGAPLPEC	332
DB	299	GQIPVGKRLQSFDEYFYFNRRCLGAPLPSC	329

RESULT 2  
T10263  
probable polygalacturonase-inhibiting protein - sweet orange  
C:Species: Citrus sinensis (sweet orange)





QY 269 -YLDLNHNQISGS-----ISSELAKD----- 269  
 Db 371 FYIDLSENEISGSLTWFFENLAHNLFEQASGNKLFEDMGKLNLSERLESIDLSENLIIFGK 430  
 QY 290 -----LQFFNVSDNNLCKKIPFGNGLORFDRFAYLHNSCICGAPLPEC 332  
 Db 431 VPMVATKQKLNLSHNLGCKLP---VTKFPASAFVGNDCGCSPLSPC 476  
 RESULT 8  
 T49908  
 hypothetical protein T24H18.110 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
 C:Accession: T49908  
 R:Bevan, M.; Robben, J.; Gyomai, B.; Volckaert, G.; Bancroft, I.; Mewes, H.W.; Rudd,  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: 225024  
 A:Accession: T49908  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-371 <BEV>  
 A:Cross-references: EMBL:AL353013; GSPDB:GNO0063; ATSP:T24H18.110  
 C:Genetics: A:Experimental source: cultivar Columbia; BAC clone T24H18  
 A:Gene: ATSP:T24H18.110  
 A:Map position: 5  
 Query Match 23.3%; Score 408.5; DB 2; Length 371;  
 Best Local Similarity 30.2%; Pred. No. 8.8e-26;  
 Matches 113; Conservative 55; Mismatches 159; Indels 47; Gaps 8;  
 QY 1 MNESFPCPLICMIFCLPNLASQRCNNNDKQALLQIKTALKNPTI--TDSWSDDD 58  
 Db 1 MDKKSLLVILLTNVVFLLSTVHS--CLPSDRAALFEKRAKNEFYIGVFTWKGCLDC 58  
 QY 59 CCGMDLVECOETSNRIISLIIDDE-----ALTGQIPROYGDLPIYQALMFK 106  
 Db 59 CKWYGVSCDPMNRVRVAGITLRGSEDDPLFOKAKRSGLMGSISPICKLTRLISGLIIAD 118  
 QY 107 LPLNFGKIPPEISALKDLKSLRSLSTSLSGVPPLFPOLTKLCLDSFKKLGIVIPOL 166  
 Db 119 WKGISVITPCIEHLPLRLHLDLVGNKFSGVITANICKLRLVNLADNHLGVIPPSI 178  
 QY 167 STPLNKAHLERNELTGEIPDIFGNFAGSPDIYLSHNLGTFVPTKTFAR---ADPRL 222  
 Db 179 TRVLSLHLDLRNNNISGVIPRDIGRKMYSRVLISGKISGQIPDSLTRIYRLAD---L 235  
 QY 223 DFGNRLREGDISLFLFGKRRKLEMLDFSGNLS-----FNFSRVO 261  
 Db 236 ELSNNRLTGPSPASFGMSVATLNDGNLISGMIPGSLASSISNLSGNLITGSIPI 295  
 QY 262 EFPSP--LTLVDLNHNQISGSLSEL-AKLDLOTFNVSDNNLCKKIPFGNGLORFDRFAY 318  
 Db 296 TFGFRSYFTVLDLANNLQGPPIASTAASFIHLIDVSHNHLGCKKIPMGSPFDHLATSP 355  
 QY 319 LHNSCICGAPLPEC 332  
 Db 356 AYNAICLCGKPLGNC 369  
 RESULT 9  
 T00971  
 Probable disease resistance protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Accession: T00971; G86459  
 C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001  
 R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
 submitted to the EMBL Data Library, April 1998  
 A:Description: Arabidopsis thaliana chromosome II BAC T9J22 genomic sequence.  
 A:Reference number: 214161

A:Accession: T00971  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-480 <R0>  
 A:Cross-references: EMBL:AC002505; NID:g2739359; PID:g2739389  
 A:Experimental source: cultivar Columbia  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umeyam, L.; Tallon,  
 euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487  
 A:Accession: G86459  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-480 <STO>  
 A:Cross-references: GB:AE002093; NID:g2739389; PID:AA014512.1; GSPDB:GNO0139  
 C:Genetics: A:Gene: T9J22.5; AT2g26380  
 A:Map position: 2  
 Query Match 22.8%; Score 401; DB 2; Length 480;  
 Best Local Similarity 26.0%; Pred. No. 5e-25;  
 Matches 123; Conservative 65; Mismatches 121; Indels 164; Gaps 17;  
 QY 15 MFLCLPNLASQRCNNNDKQALLQIKTAL-KNPT-ITDSWSDDDCGMDLVECOETSN 72  
 Db 15 VIFRLCLNPTAATCHPDEAGLAFKSGITKDPGSLSTWKKGTCCSMNGVSC-PNCG 73  
 QY 73 RIISLI--TODDA---LTGQIPROYGDLPIYQALMFKRLNPLFGKIPPEISALDKSL 127  
 Db 74 RYVVLTRIESDDAGIFLGTISPSLAKLQHLGCVFNLKNTITGPFPLFRPLRLKYV 133  
 QY 128 RUSSTLSGVPPLFPOLTKLCLDSFKKLGIVIPQS-----TLP 170  
 Db 134 YLENTLSGVLPLNICALNRLDITLYKGNRFISIPSSISNLRNLVNLGNNLTGTIP 193  
 QY 171 ---NLKA---LHLERNELTGEIPDI----- 190  
 Db 194 LGIANLKLSNLTNDGNRLSGTIPDIFKSMTNLRLITLSRNFSGKLPSIASLAPVLA 253  
 QY 191 ----GNFGSPDIYLSH-----NOLTFVPTKTFAR-----ADP 219  
 Db 254 LELGQNNLSGSPSYLSRVALDPLDLSKNRFSGAVPKSLAKLTKIANINLSHNLITNPF 313  
 QY 220 ----- 220  
 Db 314 PVLNKNYILITDLSYNKFMETIPEWTSASILGSLAKGIGKMSLDDMKTRQTDLV 373  
 QY 221 RUDSGNRLGEG-ISFLFGPKRRKLEMLDFSGNLSFNFRVDFEPPSLTYLDLNHNQISG 279  
 Db 374 SIDLSNEISGSPRLFKG-ADQLREFRMSGKRLRDLKRL-SFSTTLTDLDSRLVVG 431  
 QY 280 SUSSELAKLDTOTFNVSDNNLCKKIPFGNGLORFDRFAYLHNSCICGAPLPEC 332  
 Db 432 KVPARYA--GLKTLNLSQNHLCGKLP---VTKFPESYAGNDGCSPLSGC 478  
 RESULT 10  
 G86459  
 hypothetical protein AG26075.1 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: G86459  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo







receptor-like protein kinase - *Arabidopsis thaliana*

N;Alternate names: protein T28J14.120

C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Jan-2000 #seeds: 30 #out: 30-10-2000

C:\Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:\Accession: T18489

C; Accession: T48489  
P. Bryan M. : Murphv

R.; Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le

submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24493

A:Reference number: Z24493  
A:Accession: T48489

A;Accession: 148489  
A;Status: preliminary

```
A;status: preliminary
A;molecule type: DNA
```

A;Residues: 1-932 <BEV>

A; Cross-references: EMBL:AL163652

A; Experimental source: cultivar C

C;Genetics:

A; Map posit

A: Introns: 43/2; 67/2; 91/2; 115/2; 139/2; 163/2; 187/2; 211/2; 234/2; 258/2; 282/2; 306/2; 330/2; 354/2; 378/2; 402/2; 426/2; 450/2; 474/2; 498/2; 522/2; 546/2; 570/2; 594/2; 618/2; 642/2; 666/2; 690/2; 714/2; 738/2; 762/2; 786/2; 810/2; 834/2; 858/2; 882/2; 906/2; 930/2; 954/2; 978/2; 1002/2; 1026/2; 1050/2; 1074/2; 1098/2; 1122/2; 1146/2; 1170/2; 1194/2; 1218/2; 1242/2; 1266/2; 1290/2; 1314/2; 1338/2; 1362/2; 1386/2; 1410/2; 1434/2; 1458/2; 1482/2; 1506/2; 1530/2; 1554/2; 1578/2; 1602/2; 1626/2; 1650/2; 1674/2; 1698/2; 1722/2; 1746/2; 1770/2; 1794/2; 1818/2; 1842/2; 1866/2; 1890/2; 1914/2; 1938/2; 1962/2; 1986/2; 2010/2; 2034/2; 2058/2; 2082/2; 2106/2; 2130/2; 2154/2; 2178/2; 2202/2; 2226/2; 2250/2; 2274/2; 2298/2; 2322/2; 2346/2; 2370/2; 2394/2; 2418/2; 2442/2; 2466/2; 2490/2; 2514/2; 2538/2; 2562/2; 2586/2; 2610/2; 2634/2; 2658/2; 2682/2; 2706/2; 2730/2; 2754/2; 2778/2; 2802/2; 2826/2; 2850/2; 2874/2; 2898/2; 2922/2; 2946/2; 2970/2; 2994/2; 3018/2; 3042/2; 3066/2; 3090/2; 3114/2; 3138/2; 3162/2; 3186/2; 3210/2; 3234/2; 3258/2; 3282/2; 3306/2; 3330/2; 3354/2; 3378/2; 3402/2; 3426/2; 3450/2; 3474/2; 3498/2; 3522/2; 3546/2; 3570/2; 3594/2; 3618/2; 3642/2; 3666/2; 3690/2; 3714/2; 3738/2; 3762/2; 3786/2; 3810/2; 3834/2; 3858/2; 3882/2; 3906/2; 3930/2; 3954/2; 3978/2; 4002/2; 4026/2; 4050/2; 4074/2; 4098/2; 4122/2; 4146/2; 4170/2; 4194/2; 4218/2; 4242/2; 4266/2; 4290/2; 4314/2; 4338/2; 4362/2; 4386/2; 4410/2; 4434/2; 4458/2; 4482/2; 4506/2; 4530/2; 4554/2; 4578/2; 4602/2; 4626/2; 4650/2; 4674/2; 4698/2; 4722/2; 4746/2; 4770/2; 4794/2; 4818/2; 4842/2; 4866/2; 4890/2; 4914/2; 4938/2; 4962/2; 4986/2; 5010/2; 5034/2; 5058/2; 5082/2; 5106/2; 5130/2; 5154/2; 5178/2; 5202/2; 5226/2; 5250/2; 5274/2; 5298/2; 5322/2; 5346/2; 5370/2; 5394/2; 5418/2; 5442/2; 5466/2; 5490/2; 5514/2; 5538/2; 5562/2; 5586/2; 5610/2; 5634/2; 5658/2; 5682/2; 5706/2; 5730/2; 5754/2; 5778/2; 5802/2; 5826/2; 5850/2; 5874/2; 5898/2; 5922/2; 5946/2; 5970/2; 5994/2; 6018/2; 6042/2; 6066/2; 6090/2; 6114/2; 6138/2; 6162/2; 6186/2; 6210/2; 6234/2; 6258/2; 6282/2; 6306/2; 6330/2; 6354/2; 6378/2; 6402/2; 6426/2; 6450/2; 6474/2; 6498/2; 6522/2; 6546/2; 6570/2; 6594/2; 6618/2; 6642/2; 6666/2; 6690/2; 6714/2; 6738/2; 6762/2; 6786/2; 6810/2; 6834/2; 6858/2; 6882/2; 6906/2; 6930/2; 6954/2; 6978/2; 7002/2; 7026/2; 7050/2; 7074/2; 7098/2; 7122/2; 7146/2; 7170/2; 7194/2; 7218/2; 7242/2; 7266/2; 7290/2; 7314/2; 7338/2; 7362/2; 7386/2; 7410/2; 7434/2; 7458/2; 7482/2; 7506/2; 7530/2; 7554/2; 7578/2; 7602/2; 7626/2; 7650/2; 7674/2; 7698/2; 7722/2; 7746/2; 7770/2; 7794/2; 7818/2; 7842/2; 7866/2; 7890/2; 7914/2; 7938/2; 7962/2; 7986/2; 8010/2; 8034/2; 8058/2; 8082/2; 8106/2; 8130/2; 8154/2; 8178/2; 8202/2; 8226/2; 8250/2; 8274/2; 8298/2; 8322/2; 8346/2; 8370/2; 8394/2; 8418/2; 8442/2; 8466/2; 8490/2; 8514/2; 8538/2; 8562/2; 8586/2; 8610/2; 8634/2; 8658/2; 8682/2; 8706/2; 8730/2; 8754/2; 8778/2; 8802/2; 8826/2; 8850/2; 8874/2; 8898/2; 8922/2; 8946/2; 8970/2; 8994/2; 9018/2; 9042/2; 9066/2; 9090/2; 9114/2; 9138/2; 9162/2; 9186/2; 9210/2; 9234/2; 9258/2; 9282/2; 9306/2; 9330/2; 9354/2; 9378/2; 9402/2; 9426/2; 9450/2; 9474/2; 9498/2; 9522/2; 9546/2; 9570/2; 9594/2; 9618/2; 9642/2; 9666/2; 9690/2; 9714/2; 9738/2; 9762/2; 9786/2; 9810/2; 9834/2; 9858/2; 9882/2; 9906/2; 9930/2; 9954/2; 9978/2; 10002/2; 10026/2; 10050/2; 10074/2; 10098/2; 10122/2; 10146/2; 10170/2; 10194/2; 10218/2; 10242/2; 10266/2; 10290/2; 10314/2; 10338/2; 10362/2; 10386/2; 10410/2; 10434/2; 10458/2; 10482/2; 10506/2; 10530/2; 10554/2; 10578/2; 10602/2; 10626/2; 10650/2; 10674/2; 10698/2; 10722/2; 10746/2; 10770/2; 10794/2; 10818/2; 10842/2; 10866/2; 10890/2; 10914/2; 10938/2; 10962/2; 10986/2; 11010/2; 11034/2; 11058/2; 11082/2; 11106/2; 11130/2; 11154/2; 11178/2; 11202/2; 11226/2; 11250/2; 11274/2; 11298/2; 11322/2; 11346/2; 11370/2; 11394/2; 11418/2; 11442/2; 11466/2; 11490/2; 11514/2; 11538/2; 11562/2; 11586/2; 11610/2; 11634/2; 11658/2; 11682/2; 11706/2; 11730/2; 11754/2; 11778/2; 11802/2; 11826/2; 11850/2; 11874/2; 11898/2; 11922/2; 11946/2; 11970/2; 11994/2; 12018/2; 12042/2; 12066/2; 12090/2; 12114/2; 12138/2; 12

A;Note: T28J14.120

Query Match	19.8%;	Score 348.5;	DB 2;	Length 932;
Best Local Similarity	29.3%;	Pred. No. 2.3e-20;		
Matches 109;	Conservative 52;	Mismatches 126;	Indels 85;	Gaps 13;

QY 38 LQIKTALKN-PTITDSW--VSDDDCCGWDLVECEDETSNRILSLIQDDEALTGQIPQVG 94

Db 1 MAIKASFSNVANMLLDWDVDVHNHDFCSWRCVFCDNVSLNVVSLNLSLN-LGGEISSALG 59

QY 95 DLPYLQALWFR-----KLPN-----LFGKIPEEISALKDKLSLRSS 131

Db 60 DLMLQSIDLGKNGQIPDEIGNCVSLAYVDFSTNLLFGDIPFSISKLKQLEFLNKN 119

132 TSLSGVPLFFPQLTKLTCLDLSFNKLLGVIP-----PQLS 167

Db 120 NQLTGPIPATLTQIPNLKTLDLARNQLTGEIPRLLYWNEVLQYLGLRGMLTGTLS PDMC 179

168 TLPNKAHLERNELTGEIPDIFGNFAGSPDIYLSHNQLTGVPKTFARADPIRLDESGN 227

Db 180 QLTGLWYFDVRGNNLTGTIPESIGNCTSFIELDVSYNQITGVIPYNIQFLQVATLSLOGN 239

QY 228 RLEGDISFLFGPKRLEMLDFSGNVL-----SFNFS-----RVQFFPPS--- 266

Db 240 KLTGRIPEVIGLMQALAVLDSNLTGPPIPLGNLSFTGKLYLHGKLTGQIPPELGN 299

267 ---LTYLDLNHNQISGSLSELAKLD-LQTFNVSDNNLCGKIPTG---GNLQRFDRYAY 318

Db 300 MSRLSYQLNDNELVGKIPPELGKLEOLFELNLANNNLVGLIPSNISSCAALNQFN--- 355

319 LHN SCLCGA-PL 329 QY

Db 356 VHGNTLSGAVPL 367

Search completed: March 13, 2002, 20:57:42  
Job time: 2493 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2002, 20:56:05 ; Search time 23.91 Seconds

(without alignments)  
509.106 Million cell updates/sec

Title: US-09-308-140-7

Sequence: 1 MNEISSRCPILICIMIFCL.....FDRAYLHNSCLGAPLPEC 332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 3664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	884.5	50.4	330	1	PGIP_PYRCO
2	622	35.4	342	1	PGIP_PHAVU
3	311.5	17.7	999	1	RLK5_ARATH
4	273	15.5	980	1	CLVL_ARATH
5	236	13.4	674	1	TMIL_ARATH
6	227	12.9	942	1	TKML_ARATH
7	192	10.9	395	1	DL00_ARATH
8	185.5	10.6	1256	1	FLIT_DROME
9	175.5	10.0	2145	1	CYAA_PODAN
10	172	9.8	582	1	SHO2_MOUSE
11	169	9.6	1839	1	CYAA_SACKL
12	164.5	9.4	2300	1	CYAA_NEUCR
13	163	9.3	782	1	CHAO_TRICA
14	160.5	9.1	476	1	GP46_LEIAM
15	158	9.0	376	1	FMOD_MOUSE
16	158	9.0	582	1	SHO2_HUMAN
17	157.5	9.0	343	1	LUM_CHICK
18	157.5	9.0	713	1	GACI_HUMAN
19	154.5	8.8	375	1	FMOD_BOVIN
20	153.5	8.7	603	1	ALUS_MOUSE
21	150	8.5	376	1	FMOD_RAT
22	149.5	8.5	605	1	ALUS_PAPHA
23	147	8.4	359	1	PGS2_HUMAN
24	146.5	8.3	605	1	ALUS_HUMAN
25	146.5	8.3	2493	1	CYAA_HUMAN
26	145	8.3	376	1	FMOD_HUMAN
27	144.5	8.2	603	1	ALUS_RAT
28	143.5	8.2	338	1	LUM_HUMAN
29	143	8.1	382	1	PARC_HUMAN
30	142.5	8.1	1315	1	CHAO_DROME
31	142	8.1	331	1	PLIB_AGRAL
32	142	8.1	536	1	CBP8_HUMAN
33	142	8.1	560	1	GPV_HUMAN

34	141	8.0	277	1	RSU1_HUMAN	Q15404 homo sapien
35	138.5	7.9	312	1	AZGL_HUMAN	P02750 homo sapien
36	138.5	7.9	567	1	GPV_MOUSE	O08742 mus musculu
37	138.5	7.9	1257	1	FLIH_MOUSE	P34268 caenorhabdi
38	137	7.8	380	1	FMOD_CHICK	P51887 gallus galli
39	137	7.8	2026	1	CYAA_YEAST	P08678 saccharomyc
40	136	7.7	362	1	PGS2_PIG	O9XSD9 sus scrofa
41	135.5	7.7	1097	1	TOLL_DROME	P08993 drosophila
42	135	7.7	360	1	PGS2_BOVIN	P21793 bos taurus
43	132	7.5	354	1	PGS2_MOUSE	P28654 mus musculu
44	131.5	7.5	357	1	PGS2_CHICK	P28675 gallus galli
45	131	7.5	277	1	RSU1_MOUSE	Q01730 mus musculu

## ALIGNMENTS

```

RESULT 1
PGIP_PYRCO STANDARD: PRT: 330 AA.
AC 005091.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE POLYGALACTURONASE INHIBITOR PRECURSOR (POLYGALACTURONASE-INHIBITING
  PROTEIN).
GN PGIP.
OS Pyrus communis (Pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids I; Rosales; Rosaceae; Pyrus.
OX NCBI_TaxID=23211.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BARTLETT.
RX MEDLINE=94151422; PubMed=8108494;
RA Stotz H.U., Powell A.L., Damon S.E., Greve L.C., Bennett A.B.,
RA Labavitch J.M.;
RT "Molecular characterization of a polygalacturonase inhibitor from
RT Pyrus communis L. cv Bartlett.";
RT Plant Physiol. 102:133-138(1993).
RN [2]
RC CHARACTERIZATION.
RP STRAIN=CV. BARTLETT; TISSUE=fruit;
RT Abu-Goukh A.A., Greve L.C., Labavitch J.M.;
RT "Purification and partial characterization of 'Bartlett' pear
RT polygalacturonase inhibitors.";
RT Physiol. Mol. Plant Pathol. 23:111-122(1983).
CC -!- FUNCTION: INHIBITOR OF FUNGAL POLYGALACTURONASE. IT IS AN
CC IMPORTANT FACTOR FOR PLANT RESISTANCE TO PHYTOPATHOGENIC FUNGI.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
CC -!- SIMILARITY: TO P.VULGARIS POLYGALACTURONASE INHIBITOR.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: L09364; AAA33865.1; -
DR PIR: J02262; J02262.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR Pfam: PF00560; LRR; 4.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00370; LRR; 5.
KW Signal.
FT SIGNAL.
FT CHAIN.
FT SEQUENCE.
330 AA; 36504 MW; 36BA2BCF420393E3 CRC64;

```



"Receptor-like protein kinase genes of Arabidopsis thaliana."  
 RT Plant J. 3:451-456(1993).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV, COLUMBIA:  
 RX MEDLINE=20083488; PubMed=10617198;  
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,  
 RA Poll T., Dueterhoeft A., Stiekema W., Entian K.-D., Terryn N.,  
 RA Harris B., Anstorge W., Brandt P., Grivell L.A., Rieger M.,  
 RA Weichselgartner M., de Simone V., Obermaier B., Macho R., Mueller M.,  
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtke T.,  
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
 RA Vos P., Hohenseil J., Zimmermann M., Wedler H., Ridley P.,  
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,  
 RA Van der Schueren J., Glymoprez B., Chang Y.-J., Vandenbussche F.,  
 RA Braeken M., Weljens I., Voet M., Bastlaens I., Aert R., Deleor E.,  
 RA Wetzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,  
 RA Holzer E., Brandt A., Peters S., van Slaveren M., Dirkse W.,  
 RA Moollan P., Klein Lankhorst R., Rose M., Hauf J., Kletter P.,  
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
 RA De Keyser A., Bysschaert C., Gielen J., Villarroel R., De Clercq R.,  
 RA Van Montagu M., Rogers J., Cronin A., Quail A., Bray-Allen S.,  
 RA Clark L., Doggett J., Hall S., Kay M., Leonard N., McLay K., Mayes R.,  
 RA Petlett A., Rajandream M.-A., Lyne M., Benes V., Rechmann S.,  
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,  
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Faltmann B., Granderath K., Dauner D., Herzl A.,  
 RA Neumann S., Argitell A., Vitale D., Liguori R., Piravandi E.,  
 RA Massenot O., Quigley F., Clabaud G., Wuendlein A., Felber R.,  
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
 RA Chedford F., Cooke R., Berger C., Monfort A., Casacuberta E.,  
 RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,  
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,  
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bielek C.,  
 RA Frisman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
 RA Parnell L., Dedha N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
 RA Sekhon M., Murray J., Sheel P., Cordes M., Abu-Threideh J.,  
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,  
 RA Lattelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
 RA Mlyn P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
 RA Krimer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,  
 RA Nelson J., Spiech J., Ryan E., Andrews S., Gessel C., Layman D.,  
 RA Du H., All J., Berhoff A., Jones K., Cotton M., Joshi C.,  
 RA Antoniou B., Zidanis M., Strong C., Sun H., Lamar B., Jordan C.,  
 RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Matero A., Shah R.,  
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,  
 RA Grenat S., Shohdy N., Hasegawa A., Hameed A., Lochl M., Johnson A.,  
 RA Chen E., Marra M., Martensen R., McCombie W.R.,  
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
 thaliana."  
 RL Nature 402:769-777(1999).  
 RN [13]  
 RP CHARACTERIZATION:  
 RX MEDLINE=94368830; PubMed=8086440;  
 RA Horn M.A., Walker J.C.,  
 RT "Biochemical properties of the autophosphorylation of RLK5, a  
 receptor-like protein kinase from Arabidopsis thaliana."  
 RL Biochim. Biophys. Acta 1208:65-74(1994).  
 CC -1- COFACTOR: HAVE SIGNIFICANTLY GREATER ACTIVITY IN THE PRESENCE OF  
 MN2+ THAN MG2+.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ROOTS AND ROSETTES.  
 CC -1- PTM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.  
 CC -1- SIMILARITY: IN THE C-TERMINAL REGION; BELONGS TO THE SER/THR  
 CC FAMILY OF PROTEIN KINASES.  
 CC -1- SIMILARITY: CONTAINS 18 LEUCINE-RICH REPEATS (LRR).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>)

or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch).  
 CC -----  
 CC EMBL: M84660; AAA32859.1;  
 DR EMBL: AL021749; CAA16889.1;  
 DR EMBL: AL015722; CAB79651.1;  
 DR HSSP: P00523; 2PTK.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003592; LRR-out.  
 DR InterPro: IPR002290; Ser\_Chr\_kin\_actsite.  
 DR Pfam: PF00560; LRR\_14.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00019; LEURICRPT.  
 DR SMART: SM00370; LRR; 17.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding;  
 KW Transmembrane; Glycoprotein; Phosphorylation; Leucine-rich repeat;  
 KW Repeat; Signal.  
 FT SIGNAL 1 14  
 FT CHAIN 15 999  
 FT DOMAIN 15 621  
 FT TRANSMEM 622 641  
 FT DOMAIN 642 999  
 FT REPEAT 88 112  
 FT REPEAT 114 138  
 FT REPEAT 138 161  
 FT REPEAT 163 186  
 FT REPEAT 188 211  
 FT REPEAT 235 261  
 FT REPEAT 263 283  
 FT REPEAT 283 306  
 FT REPEAT 306 330  
 FT REPEAT 332 353  
 FT REPEAT 354 378  
 FT REPEAT 402 426  
 FT REPEAT 426 450  
 FT REPEAT 452 474  
 FT REPEAT 498 522  
 FT REPEAT 524 547  
 FT REPEAT 549 567  
 FT REPEAT 567 592  
 FT DOMAIN 683 968  
 FT CARBOHYD 98 968  
 FT CARBOHYD 102 102  
 FT CARBOHYD 150 150  
 FT CARBOHYD 185 185  
 FT CARBOHYD 210 210  
 FT CARBOHYD 259 259  
 FT CARBOHYD 282 282  
 FT CARBOHYD 452 452  
 FT CARBOHYD 576 576  
 FT NP\_BIND 689 697  
 FT BINDING 711 711  
 FT ACT\_SITE 819 819  
 FT MUTAGEN 711 711  
 SO SEQUENCE 999 AA; 109095 MW; F5793D8999EA0C6A7 CRC64;  
 K->E: LOSS OF CATALYTIC ACTIVITY.  
 K->E: LOSS OF CATALYTIC ACTIVITY.  
 Query Match 17.7%; Score 311.5; DB 1; Length 999;  
 Best Local Similarity 26.6%; Pred. No. 1,1e-17;  
 Matches 115; Conservative 49; Mismatches 125; Indels 143; Gaps 16;  
 Oy 7 FCPILICIMIFLCPNLSASORCNNDKQALQIKTKLNPTIT-DSVWSDDC--CGMD 63  
 Db 3 YCLLLCLLSLTVLPISL-----NDATILRQAKKLGLSDPAQSLSSSDNDVTPCKWL 57  
 Oy 64 LVEDEFTSN-----RII 75  
 Db 58 GVSQDASINNVSVSLSSFMVLGPPPSILCHLPJLSLSLNNYSINGSLADDFDTCHNLI 117  
 Oy 76 SLIIIDDEALIGQIRPQVG-DLPRYLQAL-----WFRKLP----- 108

```

Db 118 SLDISEN-LVGSIFPKSLPFLPMLKFLFISGNNLSDTIPSEFGEFRKLESLNAGNLS 176
Qy 109 -----NLFF-GKIPETISALKDLKSLRSLSTSGPVLPFPOLT 146
Db 177 GTPASLGWNTTLMELKAYNLFPSPQIPSPQLGNTLQVLAAGCNLYGSLPSSLT 236
Qy 147 KLTCLDLSFNKLLGVIPOLSTLPNLKALHLERNEELGCEIDIFGN-----FAGS----- 196
Db 237 SLVMDLTFNOLTSIPSWITQLTKEVQIELFNNSFGSELPSMGNNMTLKRFDASMKL 296
Qy 197 -----PD-----IYSHNOLGTFVKKTFARADPI-RLDFSGNRLEGDSIFLFGPKRK 242
Db 297 TGKIPDLNLLNLSLESLNLFENMLGEPLESTIRSKTJSELKLFNNRLTGVLPQOLGANSP 356
Qy 243 LEMDFSGNVLSPFNRSRQEPFPPS-----LTYLDLNHNQISGSLSESLAKL-DLQTFNV 295
Db 357 LQYVDLSNRRS-----GELPANCCEGCKLEYLLILDNSFGSLSNLGGCKSLTRVRL 410
Qy 296 SDNNLGGKIPFG 307
Db 411 SNRKLSCGIPHG 422

RESULT 4
CLV1_ARATH STANDARD; PRT; 980 AA.
ID CLV1_ARATH 09SY08; 004380; 09LOT2;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE RECEPTOR PROTEIN KINASE CLAVATA1 PRECURSOR (EC 2.7.1.-).
GN CLV1 OR AT1G75820 OR T4012.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG ERECTA;
RC MEDLINE=97304386; Pubmed=9160749;
RA Clark S.E., Williams R.W., Meyerowitz E.M.;
RT "The CLAVATA1 gene encodes a putative receptor kinase that controls
RT shoot and floral meristem size in Arabidopsis.";
RL Cell 89:575-585(1997).
[2]
RP SEQUENCE FROM N.A.
RP MEDLINE=99178804; Pubmed=10080719;
RA Williams R.W., Clark S.E., Meyerowitz E.M.;
RT "Genetic and physical characterization of a region of Arabidopsis
RT chromosome 1 containing the CLAVATA1 gene.";
RL Plant Mol. Biol. 39:171-176(1999).
[3]
RP SEQUENCE FROM N.A.
RP STRAIN=CV. COLUMBIA;
RC MEDLINE=21016719; Pubmed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremensskaia I., Kultz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miliuscher J., Miranda M., Nguyen M., Niemman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utechtack T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;

```

```

RT *sequence and analysis of chromosome 1 of the plant Arabidopsis
RL thaliana.";
RT Nature 408:816-820(2000).
CC -1- FUNCTION: ACTS WITH CLV3 TO CONTROL THE BALANCE BETWEEN MERISTEM
CC CELL PROLIFERATION AND DIFFERENTIATION. MAY ACT WITH CLV3 AS A
CC LIGAND-RECEPTOR PAIR IN A SIGNAL TRANSDUCTION PATHWAY.
CC COORDINATING GROWTH BETWEEN ADJACENT MERISTEMATIC REGIONS.
CC -1- SUBUNIT: MULTIMER (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC -1- TISSUE SPECIFICITY: IN A CENTRAL REGION OF THE SHOOT AND IN EARLY
CC FLOWER MERISTEMS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 18 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U96879; AAB58929.1; -.
CC EMBL: AF049870; AAD02501.1; -.
CC EMBL: AC007396; AAP26772.1; ALU_INIT.
CC Mendel: 13841; Arabid.1197;13841.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR003592; LRR_out.
CC InterPro: IPR002290; Ser_thr_kin_actsite.
CC InterPro: IPR001245; Tyr_kin.
CC Pfam: PF00569; LRR_11.
CC Pfam: PF00659; pkinase_1.
CC PRINTS: PR00019; LEURICHRPT.
CC SMART: SM00370; LRR_18.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
CC PROSITE: PS00110; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00111; PROTEIN_KINASE_ATP; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC Phosphorylation; Transmembrane.
CC
CC CHAIN 1 24
CC DOMAIN 25 980 RECEPTOR PROTEIN KINASE CLAVATA1.
CC TRANSMEM 639 659 EXTRACELLULAR (POTENTIAL).
CC DOMAIN 660 980 POTENTIAL.
CC REPEAT 93 119 CYTOPLASMIC (POTENTIAL).
CC REPEAT 143 167 LRR 1.
CC REPEAT 167 193 LRR 2.
CC REPEAT 240 264 LRR 3.
CC REPEAT 264 288 LRR 4.
CC REPEAT 288 312 LRR 5.
CC REPEAT 314 336 LRR 6.
CC REPEAT 336 360 LRR 7.
CC REPEAT 360 384 LRR 8.
CC REPEAT 384 408 LRR 9.
CC REPEAT 408 432 LRR 10.
CC REPEAT 432 455 LRR 11.
CC REPEAT 455 479 LRR 12.
CC REPEAT 479 503 LRR 13.
CC REPEAT 503 527 LRR 14.
CC REPEAT 527 551 LRR 15.
CC REPEAT 551 575 LRR 16.
CC REPEAT 575 600 LRR 17.
CC REPEAT 600 624 LRR 18.
CC REPEAT 624 648 LRR 19.
CC NP_BIND 648 706 ATP (BY SIMILARITY).
CC BINDING 706 720 ATP (BY SIMILARITY).
CC ACT_SITE 817 877 BY SIMILARITY.
CC CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 348 348 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 510 510 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 558 558 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 574 574 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 235 235 R -> P (IN REF. 3).
FT CONFLICT 703 703 A -> S (IN REF. 2).
FT CONFLICT 856 856 G -> D (IN REF. 2).
SQ SEQUENCE 980 AA; 107656 MW; B69BD7A7A8863295 CRC64;

```

Query Match 15.5%; Score 273; DB 1; Length 980;

Best Local Similarity 24.3%; Pred. No. 1,5e-14; Mismatches 105; Indels 122; Gaps 8;

```

85 LTGQIPPOVGDLPYLOALMPKLPNLFGRIPETISALKDKLSLSTSLSGVPLFPQ 144
253 LTGEITPTSLNKLHHTL-FLHINNLTGHIPEPLSGVLSKSLDLSINQLTGEIPQSF 311
145 LTKLTCGLDLSFNKLCVIPPOLSTLPNLKALHLELNELTGEIPDIFGNFAGSPDIYLSHN 204
312 LGMITLINLFNNLYGOIPRPAIGELPKLEFVEWENNFITQLPANIGRNMLTKLDVSDN 371
205 QLTGVPKTFARADPR----- 221
372 HLTGLPKDLCRGEKLEMLILSNFFGPIPEBLGCKSLTKIRIVKNLNGTVPAGLFN 431
222 -----LD-----FSG-----NR 228
432 LPLVTIETLDTNFFSELPYTMGVDVLDQIYLSNNMFSGEIPPAIGFPMQLTFLDRNR 491
229 LQSDISFLFGPKRLMLDFSGN-----VLSFNSRYQ---EFP----- 264
492 FRONIFREIFELKHLRINSANNITGIPDTSISRCSTLISVLSNNRIINGELPKGINNV 551
265 PSLTYLIDLNHNQISGLSSELAHL-DLQTFNVSNNLCKGIPFGNLQRFDRATAYLAHNSC 333
552 KNLGTINISGNQLTGIPGIGNMTSLTLDLSFNDLSGRVPGGQGLFVNEFTSFGNTY 611
QY 324 LC 325
DB 612 LC 613

```

```

RESULT 5
TML1_ARATH STANDARD: PRT: 674 AA.
AC P33543;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE KINASE-LIKE PROTEIN TMKLI PRECURSOR.
GN TMKLI.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN SEQUENCE FROM N.A.
RX STRAIN=CV. COLUMBIA; TISSUE=Green siliques;
RX MEDLINE=94033320; PubMed=8219075;
RA Valon C., Smalle J., Goodman H.M., Garaudat J.;
RT "Characterization of an Arabidopsis thaliana gene (TMKLI) encoding a
RT putative transmembrane protein with an unusual kinase-like domain.";
RC Plant Mol. Biol. 23:415-421(1993).
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -I- SIMILARITY: TO PROTEIN KINASES IN THE C-TERMINUS. BUT DOES NOT

```

```

CC SEE US HAVE CONSERVED A KINASE ACTIVITY.
CC -I- SIMILARITY: CONTAINS 7 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X72863; CAA51385.1; .
CC PIR: S35397; S35397.
CC PIR: S39476; S39476.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR003592; LRR_out.
CC Pfam: PF00560; LRR; 3.
CC Pfam: PF00699; pkinase; 1.
CC SMART: SM00370; LRR; 4.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC Receptor; Transmembrane; Glycoprotein; Signal; Leucine-rich repeat;
CC Repeat.
KW Repeat.
FT SIGNAL 1 25
FT CHAIN 26 674
FT DOMAIN 26 295
FT TRANSMEM 296 323
FT DOMAIN 324 674
FT DOMAIN 373 674
FT REPEAT 98 122
FT REPEAT 122 146
FT REPEAT 147 170
FT REPEAT 172 195
FT REPEAT 198 224
FT REPEAT 226 245
FT REPEAT 245 268
FT CARBOHYD 57 57
FT CARBOHYD 90 90
FT CARBOHYD 95 95
FT CARBOHYD 110 110
FT CARBOHYD 183 183
FT CARBOHYD 195 195
FT CARBOHYD 252 252
FT CARBOHYD 257 257
FT SEQUENCE 674 AA; 73353 MW; 533D0A8D8767E253 CRC64;

```

Query Match 13.4%; Score 236; DB 1; Length 674;

Best Local Similarity 27.1%; Pred. No. 9,6e-12; Mismatches 92; Conservative 49; Mismatches 102; Indels 96; Gaps 16;

```

QY 19 CLPNLSASQRCNNNDKQALL-QIKTALKNPT---ITDSVSDVDDCCGW-----DL 64
DB 23 CGTSLSGS-----SDVKLLGLKIKSSLLQGNSESLSSWNSSVPCQMGVWVSGSP 77
QY 65 VRC-DETSRIIISLIIODEALTLGOIPVOVGDLPYLOALMPKLP--NFGKIPETISAL 121
DB 78 LQSDISFLFGPKRLMLDFSGN-----VLSFNSRYQ---EFP----- 264
QY 122 KDKSLRLSSTSLSGVPLFPQTLTKLCLDSFNKLCVIPPOLSTL-PNLKALHLERN 180
DB 123 SMQISVFLMINISLSSGIRPLELGTSLSDVDLSGNMLAGVLPSPIMNLDKLSFRHGN 182
QY 181 ELTGEIPDIFGNFAGSPDIYLSHNQLTGFVPKTFARADPRIDLFSGNRLGDISFLGPK 240
DB 183 NISGVLP-----PAL-----PNSFCG----- 199
QY 241 KRLMLDFSGNVLSTFNSVQVEPPLST-----YLDLNHNQISGLSSELAKKLDQTFN 294
DB 200 -NLYQVLDLGNKFS-----GEPFETIRPKGVKSLDLSNNFEGVLPRGLVLESLN 252
QY 295 VSDNNLCGKIPFGNLQRFDRATAYLAHNS-CLGAPLPPEC 332

```



DB 253 LSHNNFSGMLPDFGE-SKFGAESFGNSPGLCGPLKPC 290

RESULT 6

TMK1\_ARATH STANDARD: PRT: 942 AA.

AC P43298: (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE PUTATIVE RECEPTOR PROTEIN KINASE TMK1 PRECURSOR (EC 2.7.1.-).

GN TMK1 OR F15E12.4.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI\_TaxID=3702;

XX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV\_COLUMBIA.

RX MEDLINE=93076110; PubMed=1332795;

RA Chang C., Schaller G.E., Patterson S.E., Kwok S.F., Meyerowitz E.M., Blecker A.B.;

RT "The TMK1 gene from Arabidopsis codes for a protein with structural and biochemical characteristics of a receptor protein kinase."

RL Plant Cell 4:1263-1271(1992).

RM [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV\_COLUMBIA.

RA Lin X., Kaul S., Town C.D., Benito M., Greasy T.H., Haas B.J., Wu D., Maiti R., Rongning C.M., Koo H., Fujii C.Y., Uterback T.R., Barnesstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;

RL Submitted (Oct-2000) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: PROBABLE RECEPTOR.

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -1- TISSUE SPECIFICITY: FOUND IN MOST TISSUES.

CC -1- PTM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC -1- SIMILARITY: CONTAINS 9 LECITINE-RICH REPEATS (LRR).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

CC EMBL: L00670; AAA32876.1; .

DR EMBL: AC026480; AAC51302.1; .

DR HSSP: P11362.1FCI.

DR InterPro: IPR000719; Euk\_pkinase.

DR InterPro: IPR001611; LRR.

DR InterPro: IPR003592; LRR\_out.

DR InterPro: IPR003591; LRR\_Typ.

DR InterPro: IPR002290; Ser\_thr\_kin\_actsite.

DR Pfam: PF00560; LRR: 8.

DR Pfam: PF00069; pkinase: 1.

DR SMART: SM00370; LRR: 7.

DR SMART: SM00369; LRR\_Typ: 1.

DR SMART: SM00221; STYK: 1.

DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP: 1.

DR PROSITE: PS00108; PROTEIN\_KINASE\_ST: 1.

DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM: 1.

KW Transmembrane; Serine/threonine-protein kinase; ATP-binding;

KW Transmembrane; Receptor; Glycoprotein; Signal; Repeat;

KW Leucine-rich repeat; Phosphorylation.

FT SIGNAL 1 23 POTENTIAL.

FT CHAIN 24 942 PUTATIVE RECEPTOR PROTEIN KINASE TMK1.

FT DOMAIN 24 482 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 483 503 POTENTIAL.

FT DOMAIN 504 942 CYTOPLASMIC (POTENTIAL).

FT REPEAT 87 110 LRR 1.

FT REPEAT 110 133 LRR 2.

FT REPEAT 135 159 LRR 3.

FT REPEAT 185 209 LRR 4.

FT REPEAT 231 253 LRR 5.

FT REPEAT 254 278 LRR 6.

FT REPEAT 280 300 LRR 7.

FT REPEAT 385 408 LRR 8.

FT REPEAT 409 436 LRR 9.

FT DOMAIN 588 869 PROTEIN KINASE.

FT NP\_BIND 594 602 ATP (BY SIMILARITY).

FT BINDING 616 616 ATP (BY SIMILARITY).

FT ACT\_SITE 717 717 BY SIMILARITY.

FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 363 363 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 942 AA; 102387 MW; 93E30B52F5F549DE CRC64;

Query Match 12.9%; Score 227; DB 1; Length 942;

Best Local Similarity 28.8%; Pred. No. 8,1e-11;

Matches 88; Conservative 52; Mismatches 138; Indels 28; Gaps 14;

QY 11 LCICMIFCLPNLASQRCNNNDKQALLQITALKNPTITDSWVDDCCGMDVECDT 70

DB 8 LLEFSTFULL--LBSKADSDGDLASAMLSKLKSL-NPPSSFGW-SDPDCKWTHIVC--T 61

QY 71 SNRIISLIIDDEALTIQIPQVGDLPYQAL--WPKRLPNLFKIPPEISALKDKSL 127

DB 62 GTKRVTRIQIHSGLQGLTSLPDLNLSELERLEQW---NNISGVPV-SLSGLASLQVL 116

QY 128 RLSTSLSGPRLPEFPQTLTKCLDLSFNKLG--VIRPOLSTLNKALHRELHELGEI 186

DB 117 MLSNNNDSDIPSDVFOGLTSLQSVIEIDNNPKSWEIPEISLNASALQFNSANVSGSL 176

QY 187 PDIG--NFASSPDIVLSHNOITGFVPTFARADPIRLDFSGNRLGDISFL--FGPK 241

DB 177 PGCFGPEPEPFLSLHLAFNNLEBELPMSLAGSVQSLMNLGQKLTGDIYLVQMTGKE 236

QY 242 -RLMDLDSGVALSFNSRVOEFPFSLTYLDLNHNQISGLSSLELAKID-LQTNVSDNN 299

DB 237 VMLHSNKFSGPL--DFSGIKE---LESLSLRDNSFGVPASLSLESILKVVNLNNH 290

QY 300 LCGKIP 305

DB 291 LQGPVP 296

RESULT 7

D100\_ARATH STANDARD: PRT: 395 AA.

AC Q00874;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE DNA-DAMAGE-REPAIR/TOLERATION PROTEIN DRT100 PRECURSOR (FRAGMENT).

GN DRT100.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI\_TaxID=3702;

XX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV\_COLUMBIA.

RX MEDLINE=92390391; PubMed=1518832;

RA Pang O., Hays J.B., Rajagopal I.;

RT "A plant cDNA that partially complements Escherichia coli reca

RT mutations predicts a polypeptide not strongly homologous to RecA  
 RT proteins".  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:8073-8077(1992).  
 CC -i- FUNCTION: THIS PROTEIN IS ABLE TO COMPLEMENT BACTERIAL RECA  
 CC MUTATIONS, BUT IT'S NATIVE FUNCTION IN THE PLANT IS NOT KNOWN.  
 CC -i- SIMILARITY: CONTAINS 4 LEUCINE-RICH REPEATS (LRR).  
 CC -i- SIMILARITY: STRONG TO OTHER EUKARYOTIC RECA-LIKE PROTEIN; SOME, TO  
 CC PROKARYOTIC RECA PROTEIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X66482; CAA47109.2; .  
 DR PIR: A46260; A46260.  
 DR PIR: S22863; S22863.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003592; LRR\_out.  
 DR Pfam: PF00560; LRR; 1.  
 DR SMART: SM00370; LRR; 3.  
 DR Transic peptide: Chloroplast; leucine-rich repeat; Repeat:  
 KW ATP-binding; DNA repair; DNA damage.  
 KW NON\_TER 1  
 FT TRANSIT <1 73 CHLOROPLAST (POTENTIAL).  
 FT CHAIN 74 395 DNA-DAMAGE-REPAIR/TOLERATION PROTEIN  
 FT DRT100.  
 FT REPEAT 155 180 LRR 1.  
 FT REPEAT 250 274 LRR 2.  
 FT REPEAT 276 297 LRR 3.  
 FT REPEAT 298 322 LRR 4.  
 FT NP\_BIND 189 196 ATP (POTENTIAL).  
 SQ SEQUENCE 395 AA: 43810 MW: 48434.33 kDa  
 Query Match 10.9%; Score 192; DB 1; Length 395;  
 Best Local Similarity 22.4%; Pred. No. 1.9e-08;  
 Matches 83; Conservative 43; Mismatches 144; Indels 100; Gaps 8;

ID FLII\_DROME STANDARD: PRT: 1256 AA.  
 AC 024020: Q24088; Q9VRH0;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE FLIGHTLESS-I PROTEIN.  
 GN FLII OR CG1484.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OREGON-R; TISSUE=Embryo;  
 RX MEDLINE=94068608; PubMed=8248259;  
 RA Campbell H.D., Schmankey T., Claudianos C., Otsarac N.,  
 RA Kasprzak A.B., Cotseal J.N., Young I.G., de Couet H.G., Miklos G.L.G.;  
 RA "The Drosophila melanogaster flightless-I gene involved in  
 RT gastrulation and muscle degeneration encodes gelsolin-like and  
 RT leucine-rich repeat domains and is conserved in Caenorhabditis elegans  
 RT and humans".  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:11386-11390(1993).  
 [2]  
 RP SEQUENCE FROM N.A., AND VARIANT SER-601.  
 RC STRAIN=CANTON-S;  
 RX MEDLINE=96129280; PubMed=8582612;  
 RA de Couet H.G., Fong K.S.K., Weeds A.G., McLaughlin P.J.,  
 RA Miklos G.L.G.;  
 RA "Molecular and mutational analysis of a gelsolin-family member encoded  
 RT by the flightless I gene of Drosophila melanogaster".  
 RL Genetics 141:1049-1059(1995).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CANTON-S;  
 RX MEDLINE=98188272; PubMed=9520435;  
 RA Maleszka R., de Couet H.G., Miklos G.L.G.;  
 RA "Data transferability from model organisms to human beings: insights  
 RT from the functional genomics of the flightless region of Drosophila".  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:3731-3736(1998).  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY; TISSUE=Embryo;  
 RX MEDLINE=20196012; PubMed=10731138;  
 RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Fitse E.,  
 RA Stapleton M., Harvey D.A.;  
 RA "A Drosophila complementary DNA resource".  
 RL Science 287:2222-2224(2000).  
 [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champs M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshfegh A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskaas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenhach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RL "The genome sequence of *Drosophila melanogaster*."  
Science 287:2185-2195(2000).  
CC -1- FUNCTION: MAY PLAY A KEY ROLE IN EMBRYONIC CELLULARIZATION BY  
CC INTERACTING WITH BOTH THE CYTOSKELETON AND OTHER CELLULAR  
CC COMPONENTS. ALTERNATIVELY, IT MAY PLAY A STRUCTURAL ROLE IN  
CC INDIRECT FLIGHT MUSCLE VITAL FOR EMBRYONIC DEVELOPMENT.  
CC -1- TISSUE SPECIFICITY: FOUND IN OVARIES, LARVAL FAT BODIES, BRAIN AND  
CC ADULT THORAX.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY.  
CC -1- DOMAIN: CONSISTS OF A LEUCINE-RICH AMINO TERMINAL HALF, WHICH IS  
CC LIKELY TO BE INVOLVED IN PROTEIN-PROTEIN INTERACTION, AND A  
CC CARBOXYTERMINAL HALF WHICH HAS HIGH SEQUENCE SIMILARITY TO  
CC GELSOLIN AND IS THEREFORE LIKELY TO BE INVOLVED IN ACTIN-BINDING.  
CC -1- SIMILARITY: BELONGS TO THE VILLIN/GELSOLIN FAMILY.  
CC -1- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).  
CC -1- SIMILARITY: CONTAINS 4 GELSOLIN-LIKE REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U01182; AAC03566.1; -  
CC EMBL: AF017777; AAC28407.1; -  
CC EMBL: AF132184; AAD34772.1; -  
CC EMBL: AE003568; AAF50830.1; ALT\_SEQ.  
CC HSP: P02640; 2YIL  
CC FlyBase: FBgn0000709; f111.  
CC InterPro: IPR001974; Gelsolin.  
CC InterPro: IPR001611; LRR.  
CC InterPro: IPR003592; LRR\_out.  
CC InterPro: IPR003591; LRR\_typ.  
CC Pfam: PF00626; Gelsolin; 4.  
CC Pfam: PF00560; LRR; 12.  
CC PRINTS: PR00019; LEURICHRPT.  
CC PRINTS: PR00597; GELSOLIN.  
CC SMART: SM00262; GEL; 6.  
CC SMART: SM00370; LRR; 3.  
CC SMART: SM00369; LRR\_typ; 2.  
CC Developmental protein; Repeat: Leucine-rich repeat; Polymorphism.  
CC Repeat 4  
CC REPEAT 29 LRR 1.  
CC REPEAT 51 LRR 2.  
CC REPEAT 52 LRR 3.  
CC REPEAT 75 LRR 4.  
CC REPEAT 100 LRR 5.  
CC REPEAT 124 LRR 6.  
CC REPEAT 145 LRR 7.  
CC REPEAT 169 LRR 8.  
CC REPEAT 171 LRR 9.  
CC REPEAT 218 LRR 10.  
CC REPEAT 243 LRR 11.  
CC REPEAT 264 LRR 12.  
CC REPEAT 289 LRR 13.  
CC REPEAT 313 LRR 14.  
CC REPEAT 336 LRR 15.  
CC REPEAT 360 LRR 15.

FT REPEAT 499 557 GELSOLIN-LIKE 1.  
FT REPEAT 746 789 GELSOLIN-LIKE 2.  
FT REPEAT 1064 1102 GELSOLIN-LIKE 3.  
FT REPEAT 1165 1206 GELSOLIN-LIKE 4.  
FT VARIANT 601 601 G->S.  
FT CONFLICT 1067 1070 STFC->HYFS (IN REF. 5).  
FT CONFLICT 1068 1068 T->A (IN REF. 2).  
SO SEQUENCE 1256 AA; 143681 MW; CF0056EFAA8DB92 CRC64;  
  
Query Match 10.6%; Score 185.5; DB 1; Length 1256;  
Best Local Similarity 30.9%; Pred. 2.9e-07;  
Matches 81; Conservative 36; Mismatches 120; Indels 25; Gaps 11;  
  
QY 82 DEALTGIPPOVGDPLQ--ALMFRKPLNLFKGIPELISALDKSLRSLSGPPV 139  
DB 37 DRQLAIPREILHLKLEHLSLNHNLEKIFGEI-TELSCLSL-DLRNQKNGGIP 94  
QY 140 LFFPOLTKLCLDLSEFKLGLVIPPOLSTPLNKLALHRENETGEIPDFGFGSPDI 199  
DB 95 ELF-HLEELTLDLSHNKKEV-PEGLERAKNLIVLSNNOIESIPTPLFIHTDLFL 152  
QY 200 YLSHNDLTGFVPTFARADPIRLDFSGNLEGDISFLGPK-----KLEMDFGSNVLS 254  
DB 153 DLSHNRLETLPQTRRLINLKTLDLSHNPLE-----LFQLRQLPSLOSLEVLKMSGTQT 207  
QY 255 -PNSRQVEPPSLTYLDLHNOISGSLSELAKLDLOTNVSDNNLCGRPTGCVL-QR 312  
DB 208 LNLNPPSIDSLANLCEDLSHNSLPKLPDCVNVVYTLVLRISDNELT-ELTAGVELMOR 266  
QY 313 FDRAYLHN-----SCLGAP 328  
DB 267 LESLNLSRNQVALLPALCKLP 288  
  
RESULT 9  
CYAA\_PODAN  
ID CYAA\_PODAN STANDARD: PRT: 2145 AA.  
AC 001513;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ADENYLATE CYCLASE (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE) (ADENYLYL  
DE CYCLASE).  
OS Podospora anserina.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Lasiosphaeriaceae; Podospora.  
OX NCBI\_TaxID=5145;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96200667; PubMed=8621071;  
RA Loubradou G., Begueret J., Turcu B.,  
RT "An additional copy of the adenylate cyclase-encoding gene relieves  
RT developmental defects produced by a mutation in a vegetative  
RT incompatibility-controlling gene in *Podospora anserina*."  
RL Gene 170:119-123(1996).  
CC -1- FUNCTION: PLAYS ESSENTIAL ROLES IN REGULATION OF CELLULAR  
CC METABOLISM BY CATALYSING THE SYNTHESIS OF A SECOND MESSENGER,  
CC CAMP.  
CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-CYCLIC AMP + PYROPHOSPHATE.  
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-3 FAMILY.  
CC -1- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).  
CC -1- SIMILARITY: CONTAINS A P2C-LIKE DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: LA3413; AAB05642.1; -

```

DR InterPro: IPR001054; Guanylt_cyclase.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_typ.
DR InterPro: IPR003589; PP2C_catalytic.
DR InterPro: IPR001932; PP2C_domain.
DR InterPro: IPR000159; RA.
DR Pfam: PF00211; guanylate_cyc. 2.
DR Pfam: PF00360; LRR_13.
DR Pfam: PF00481; PP2C_1.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00044; CYCC_1.
DR SMART: SM00370; LRR_5.
DR SMART: SM00369; LRR_TYP_1.
DR SMART: SM00332; PP2CC_1.
DR SMART: SM00314; RA_1.
DR Lyase; Repeat; Leucine-rich repeat; cAMP synthesis; Magnesium.
KW Lyase; Repeat; Leucine-rich repeat; cAMP synthesis; Magnesium.
FT REPEAT 751 776 LRR 1.
FT REPEAT 778 800 LRR 2.
FT REPEAT 801 824 LRR 3.
FT REPEAT 826 847 LRR 4.
FT REPEAT 848 870 LRR 5.
FT REPEAT 872 894 LRR 6.
FT REPEAT 895 917 LRR 7.
FT REPEAT 919 940 LRR 8.
FT REPEAT 941 965 LRR 9.
FT REPEAT 963 1005 LRR 10.
FT REPEAT 1006 1027 LRR 11.
FT REPEAT 1028 1051 LRR 12.
FT REPEAT 1053 1074 LRR 13.
FT REPEAT 1075 1097 LRR 14.
FT REPEAT 1099 1120 LRR 15.
FT REPEAT 1123 1256 LRR 16.
FT REPEAT 1257 1280 LRR 17.
FT REPEAT 1282 1304 LRR 18.
FT REPEAT 1305 1328 LRR 19.
FT REPEAT 1357 1380 LRR 20.
FT REPEAT 1394 1419 LRR 21.
FT DOMAIN 1445 1710 PP2C-LIKE.
FT DOMAIN 1711 2145 CATALYTIC.
FT DOMAIN 36 41 POLY-SER.
SQ SEQUENCE 2145 AA; 237515 MW; 88E7E6E4AC0687D CRC64;

```

```

Query Match 10.0%; Score 175.5; DB 1; Length 2145;
Best Local Similarity 23.0%; Pred. No. 3.9e-06;
Matches 76; Conservative 56; Mismatches 111; Indels 87; Gaps 14;

```

```

OY 16 IFCLPNTLSASORCNNDKQALLQIKTALKNPITTDSDVSDCCGMDLVECDTSNRIT 75
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
DB 798 ISVC-PNL-RQIKFRNNENARALPKSPGYASRLTMDA-----SNNRLE 838
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
OY 76 SLTIQDEALTGQIPQVGDLPYLQAWFRKLPNLFGRKIPPEISALMDKSLRLSSLS 135
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
DB 839 SLSSAALHNHLGLGLKMLAN-----NKLQLPREFAPAVLKTNLSSMLN 885
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
OY 136 GPVLPFPQTLKTLCDLSENKLLGVIPOLSTLPNKLALHLENNELTGEIPDIFGNFAG 195
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
DB 886 N-FPPFLAKLENLVDLDSFNITQS-LPDVNGQWTSLERLVITNNELSGSLPSPFKLRS 943
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
OY 196 SPDIYLSHNDLGT-----VPKTFARADPIRLDPSGGRLEGDISFLGPKKRLMDLDFSG 250
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
DB 944 LRELDIYVNMISNDIVISQLPK-----LEILSATFNNTIS--QFSG 981
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
OY 251 NV-----LSFNFSRVOEFP-----PSLTVDLNNHNOISGSLSELAKLDLOTFNVS----- 296
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
DB 982 TFERVRSIKLNMNNTITFELKAPVPTLKALNTLS-----NAQLASIDSFHMSNLER 1033
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
OY 297 ----DNILCGKIPFG-GNLQRPDRTAYLANHS 322
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
DB 1034 LEDDKNFVSLPAHIGNLRLREVEFSIAHNS 1063
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |

```

```

RESULT 10
SHO2_MOUSE
ID SHO2_MOUSE STANDARD; PRT; 582 AA.
AC 088520;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LEUCINE-RICH REPEAT PROTEIN SHOC-2 (RAS-BINDING PROTEIN SUR-8).
GN SHOC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9837190; PubMed=9674433;
RA Sieburth D.S., Sun Q., Han M.;
RT "SUR-8, a conserved Ras-binding protein with leucine-rich repeats,
RT positively regulates Ras-mediated signaling in C. elegans."
RL Cell 94:119-130(1998).
CC -I SUBUNIT: SPECIFICALLY BINDS K-RAS AND N-RAS BUT NOT H-RAS IN
CC VITRO.
CC -I SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF068921; AAC40175.1; -
DR MGD; MGI:1927197; Shoc2.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF00560; LRR_16.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00370; LRR_8.
DR SMART: SM00369; LRR_TYP_5.
KW Repeat; Leucine-rich repeat.
FT REPEAT 99 122 LRR 1.
FT REPEAT 122 145 LRR 2.
FT REPEAT 145 168 LRR 3.
FT REPEAT 168 191 LRR 4.
FT REPEAT 193 214 LRR 5.
FT REPEAT 214 237 LRR 6.
FT REPEAT 239 260 LRR 7.
FT REPEAT 262 283 LRR 8.
FT REPEAT 283 306 LRR 9.
FT REPEAT 306 329 LRR 10.
FT REPEAT 331 353 LRR 11.
FT REPEAT 354 377 LRR 12.
FT REPEAT 379 400 LRR 13.
FT REPEAT 401 424 LRR 14.
FT REPEAT 424 447 LRR 15.
FT REPEAT 447 471 LRR 16.
FT REPEAT 473 493 LRR 17.
FT REPEAT 493 516 LRR 18.
FT REPEAT 518 540 LRR 19.
FT REPEAT 540 563 LRR 20.
SQ SEQUENCE 582 AA; 64947 MW; 62C5C082B3CF5F12 CRC64;

```

```

Query Match 9.8%; Score 172; DB 1; Length 582;
Best Local Similarity 26.6%; Pred. No. 1.3e-06;
Matches 76; Conservative 35; Mismatches 113; Indels 62; Gaps 12;

```

```

OY 63 DLVECDTSNRITSLTIQDEALTGQIPQVGDLPYLQAWFRKLPNLFGRKIPPEISALK 122
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |
DB 93 ELNCKRENSMRDLKSRSHIL-----PPSVKELTQLTETLY--YSNKKQSLPAVGVCLV 146
    | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | |

```

OY 123 DUKSLRSTSLGSPVLPFPOLTKLTCLDLSFNKL-----LG 160  
 DB 147 NLMTLASENSLTS-LPDSLDNLKLRMLDLRHKKLRIPSVYRLDSTLYLRFNRT 205  
 OY 161 VIPPOLSTLPNLKALHLEHRELGEIPDIFGNFAGSPDIYLSHNOITGFEVKTARADPI 220  
 DB 206 TVERDIKLPKLSMLSTIRENIK-QLPALIGELCNLTITLDAVHNOLE-HLPKEIGNCTOI 263  
 OY 221 -RLDFSGNRLEGDISFLFGPKKRLKLEMDFGSNVLSFN-----FSRVGEFPPSLT----- 268  
 DB 264 TNLDLQHNLDL-----LDLPDTIGNLSSLNKLGRNRLSAIPRSLAKCSALE 310  
 OY 269 YLDLNNHOIS---GSLSELAKLDQTFNVSNDNLCKGIPFGNQL 311  
 DB 311 ELNLBNNTISTLPESLSLVLKNSLTL---ARKCPOLYVGGSQ 353  
 RESULT 11  
 CYAA\_SACKL STANDARD: PRT: 1839 AA.  
 AC P23466;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE ADENYLATE CYCLASE (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE) (ADENYLYL CYCLASE).  
 OS Saccharomyces kluyveri (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91323718; PubMed=1864503;  
 RA Young D., O'Neill K., Broek D., Wigler M.;  
 RL "The adenylate cyclase-encoding gene from Saccharomyces kluyveri.";  
 RL Gene 102:129-133(1991).  
 CC -!- FUNCTION: PLAYS ESSENTIAL ROLES IN REGULATION OF CELLULAR  
 CC METABOLISM BY CATALYSING THE SYNTHESIS OF A SECOND MESSENGER,  
 CC CAMP.  
 CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-CYCLIC AMP + PYROPHOSPHATE.  
 CC -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-3 FAMILY.  
 CC -!- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).  
 CC -!- SIMILARITY: CONTAINS A PP2C-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X56042; CAA39513.1; .  
 DR PIR: S14464; OYBYK  
 DR PIR: J01145; J01145.  
 DR InterPro: IPR001054; Guanylyl\_cyclase.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003592; LRR\_out.  
 DR InterPro: IPR003591; LRR\_Typ.  
 DR InterPro: IPR003589; PP2C\_catalytic.  
 DR InterPro: IPR001932; PP2C\_domain.  
 DR InterPro: IPR000159; RA.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR Pfam: PF00560; LRR\_14.  
 DR Pfam: PF00481; PP2C\_1.  
 DR PRINTS: PR00019; LEURICHRPT.  
 DR SMART: SM00044; CYCC; 1.  
 DR SMART: SM00370; LRR; 4.  
 DR SMART: SM00369; LRR\_Typ; 3.  
 DR SMART: SM00352; PP2C; 1.  
 DR SMART: SM00314; RA; 1.

DR PROSITE: PSS0125; GUANYLATE\_CYCLASES\_2; 1.  
 KM Lysase; Repeat; Leucine-rich repeat; CAMP synthesis; Magnesium.  
 FT REPEAT 632 655 LRR 1.  
 FT REPEAT 659 679 LRR 2.  
 FT REPEAT 680 702 LRR 3.  
 FT REPEAT 703 724 LRR 4.  
 FT REPEAT 726 748 LRR 5.  
 FT REPEAT 749 771 LRR 6.  
 FT REPEAT 773 793 LRR 7.  
 FT REPEAT 794 818 LRR 8.  
 FT REPEAT 820 833 LRR 9.  
 FT REPEAT 834 857 LRR 10.  
 FT REPEAT 859 879 LRR 11.  
 FT REPEAT 880 903 LRR 12.  
 FT REPEAT 905 926 LRR 13.  
 FT REPEAT 927 949 LRR 14.  
 FT REPEAT 951 974 LRR 15.  
 FT REPEAT 1004 1025 LRR 16.  
 FT REPEAT 1025 1048 LRR 17.  
 FT REPEAT 1050 1072 LRR 18.  
 FT REPEAT 1074 1096 LRR 19.  
 FT REPEAT 1101 1124 LRR 20.  
 FT REPEAT 1135 1160 LRR 21.  
 FT DOMAIN 1185 1440 PP2C-LIKE.  
 FT DOMAIN 1441 1839 CATALYTIC.  
 SQ SEQUENCE 1839 AA; 206895 MW; 86A69BCB1F2733CB CRC64;

Query Match 9.6%; Score 169; DB 1; Length 1839;  
 Best Local Similarity 23.6%; Pred. No. 1; 1e-05;  
 Matches 86; Conservative 48; Mismatches 122; Indels 108; Gaps 15;

OY 44 LKNPTITDSWVS--DDDCGMDLYECDETSNRITSLT----- 78  
 DB 816 LQNLFLTDNRISTFDLDTLRRLTELQO--NPITSMVCGVYMANMTSLSNKAKLSFS 873  
 OY 79 -----IODEALTGQIPPOVDLPYLOALWFRKLPNLFGRKIPERISALKDKLS 126  
 DB 874 AELLSKLPRLKLELLENENLQLPPEINKLRL--YLSVARNNLESIPDEISDLRSLS 931  
 OY 127 LRLSS-----TSLSGVPVLPF-----POLTK-LTCLDLSFN 156  
 DB 932 LDLSNNMLRLMNNLEDELTSLVSSNLLTGFGSPAKFPASPSPKLAKSLDLFSYADN 991  
 OY 157 KLGVIIPOLSTENKALHLEHRELGEIPDIFGNFAGSPDIYLSHNOITGFEVKTAR 216  
 DB 992 NLDSIMPLVNTFQNLKTLNLSYNNFV-EISDL--KQNLTELTLGNNFTSLPGEAVQH 1048  
 OY 217 ADPIR-LDFSGNR---LEGDISFLFGPKKRLKLEMDFGSNVLSFNFSRVQ-----EPFPSL 267  
 DB 1049 LRSKLVLMLNKNKLSLPALISQL-----SRLSVLDVGSNQLKINISYHYDMNRRNKDL 1104  
 OY 268 TYLDLNNH---QISGLSSSE-----LAKLDQTFNVSNDNLCKGIPT 306  
 DB 1105 KYLNFSGNKRFEIKSALDPEDGKNLSDGLIKQLRVGLMDVTLTKSKVDSEVSIRLRT 1164  
 OY 307 GGNL 310  
 DB 1165 TASM 1168  
 RESULT 12  
 CYAA\_NEUCR STANDARD: PRT: 2300 AA.  
 AC Q01631;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE ADENYLATE CYCLASE (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE) (ADENYLYL CYCLASE).  
 GN CR-1 OR NAC.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Sordariales: Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92000795; PubMed=1680356;  
 RA Kore-Eda S., Murayama T., Uno I.;  
 RT "Isolation and characterization of the adenylate cyclase structural  
 gene of *Neurospora crassa*.";  
 RL Jpn. J. Genet. 66:317-334(1991).  
 CC -1- FUNCTION: PLAYS ESSENTIAL ROLES IN REGULATION OF CELLULAR  
 CC METABOLISM BY CATALYSING THE SYNTHESIS OF A SECOND MESSENGER,  
 CC CAMP.  
 CC -1- CATALYTIC ACTIVITY: ATP -> 3',5'-CYCLIC AMP + PHOSPHOSPHATE.  
 CC -1- SIMILARITY: BELONGS TO ADENYLATE CYCLASE CLASS-3 FAMILY.  
 CC -1- SIMILARITY: CONTAINS 23 LEUCINE-RICH REPEATS (LRR).  
 CC -1- SIMILARITY: CONTAINS A PP2C-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D00909; BAA00755.1; -  
 DR InterPro: IPR001054; Guanylate\_cyclase.  
 DR InterPro: IPR001811; LRR.  
 DR InterPro: IPR003592; LRR\_out.  
 DR InterPro: IPR003591; LRR\_tyr.  
 DR InterPro: IPR003589; PP2C\_catalytic.  
 DR InterPro: IPR001932; PP2C\_domain.  
 DR InterPro: IPR000159; RA.  
 DR Pfam: PF00211; guanylate\_cyc; 2.  
 DR Pfam: PF00560; LRR; 13.  
 DR Pfam: PF00481; PP2C; 1.  
 DR PRINTS: PR00019; LEURICHRPT.  
 DR SMART: SM00044; CYCC; 1.  
 DR SMART: SM00370; LRR; 4.  
 DR SMART: SM00369; LRR\_TYR; 2.  
 DR SMART: SM00332; PP2C; 1.  
 DR SMART: SM00314; RA; 1.  
 DR PROSITE: PS50125; GUANYLATE\_CYCLASES\_2; 1.  
 KM LRRase: Repeat: Leucine-rich repeat; CAMP synthesis; Magnesium.  
 FT REPEAT 867 890 LRR 1.  
 FT REPEAT 892 914 LRR 2.  
 FT REPEAT 915 938 LRR 3.  
 FT REPEAT 938 961 LRR 4.  
 FT REPEAT 962 986 LRR 5.  
 FT REPEAT 988 1008 LRR 6.  
 FT REPEAT 1009 1031 LRR 7.  
 FT REPEAT 1033 1055 LRR 8.  
 FT REPEAT 1056 1079 LRR 9.  
 FT REPEAT 1081 1097 LRR 10.  
 FT REPEAT 1098 1120 LRR 11.  
 FT REPEAT 1122 1142 LRR 12.  
 FT REPEAT 1143 1165 LRR 13.  
 FT REPEAT 1166 1188 LRR 14.  
 FT REPEAT 1189 1211 LRR 15.  
 FT REPEAT 1213 1234 LRR 16.  
 FT REPEAT 1349 1369 LRR 17.  
 FT REPEAT 1373 1396 LRR 18.  
 FT REPEAT 1398 1420 LRR 19.  
 FT REPEAT 1422 1445 LRR 20.  
 FT REPEAT 1474 1497 LRR 21.  
 FT REPEAT 1513 1538 LRR 22.  
 FT REPEAT 1564 1829 LRR 23.  
 FT DOMAIN 1830 2300 PP2C-LIKE.  
 FT DOMAIN 49 52 CATALYTIC.  
 FT DOMAIN 204 221 POLY-SER.  
 FT SEQUENCE 2300 AA; 254752 MW; 52E79B90E6B17A7B CRC64;

Query Match 9.4%; Score 164.5; DB 1; Length 2300;  
 Best Local Similarity 23.8%; Pred No. 3.3e-05;  
 Matches 72; Conservative 45; Mismatches 125; Indels 61; Gaps 11;  
 QY 30 NNNDKOAL-LQIKTKLNPTITDSWVSDCCGWDLVECDSTNRISLIIDDEALTSO 88  
 DB 924 NNNEQALPKSPFATASK-----LTYDVGSNNRQDDHSELKSLITL 965  
 QY 89 IPPQGGDLPYDIALMFRKLPNLFKIPETISALKDKLSRLSTSLSCGVPPLFFPOLTKL 148  
 DB 966 LKVNLAN-----NCLRSLPPLTGAYKSLRTLNISNFDL-VFSPFICELETI 1011  
 QY 149 TGLDLSFNKLCVIPPOLSTLPNLKALHLERNELGEPIDGNGFAGSPDIYLSNOLTG 208  
 DB 1012 VOLDSLFSN-INNLNDNLKMLNLEKFTYNNRSLGPSSESVRLDLSRELDIRYNQIST 1070  
 QY 209 F-VPTFRADPRIRLDFSGNRLGDISFLFGPKRLEMLDFSGN-VLSFNSRVOEFPSPS 266  
 DB 1071 IDVLSDLPLRLTSLAD-----HNQISKFSGSEFRRLSLKNSNPIVKEVYKAPVIP- 1122  
 QY 267 LTYDLNNNOISGSLSELAKDLDTFN-----VSDNNLCGKIPTG-GNLQRFDRYAY 318  
 DB 1123 -----QILNLSMAQLASIDESIDNLMLERLILDSNFFVSLPNOIGNLKKLDHLSM 1173  
 QY 319 LHN 321  
 DB 1174 ANN 1176  
 RESULT 13  
 ID CHAO-TRICA STANDARD; PRT: 782 AA.  
 AC P82963;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DE 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE CHAOPTIN (PHOTORECEPTOR CELL-SPECIFIC MEMBRANE PROTEIN) (FRAGMENT).  
 GN CHP.  
 OS Tribolium castaneum (Red flour beetle).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
 OC Cucujiformia; Tenebrionidae; Tribolium.  
 OX NCBI\_TaxID=7070;  
 RN 11;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GA-1;  
 RA Nie W., Stronach B., Panganiban G., Brown S., Denell R.;  
 RT "Molecular characterization of TcIab1 and the 3' end of the  
 RT tribolium homeotic complex.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: REQUIRED FOR PHOTORECEPTOR CELL MORPHOGENESIS. MEDIATES  
 CC HOMOPHILIC CELLULAR ADHESION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR SURFACE OF R-CELL PLASMA  
 CC MEMBRANE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE CHAOPTIN FAMILY.  
 CC -1- SIMILARITY: CONTAINS AT LEAST 21 LEUCINE-RICH REPEATS (LRR).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF322227; AAK01654.1; -  
 DR InterPro: IPR001611; LRR.  
 DR Pfam: PF00560; LRR; 29.  
 DR PRINTS: PR00019; LEURICHRPT.  
 KW Glycoprotein; Membrane; Repeat; Leucine-rich repeat; Vision.  
 FT NON\_TER 1 1  
 FT REPEAT 43 66 LRR 1.

```

FT REPEAT 67 90 LRR 2.
FT REPEAT 93 116 LRR 3.
FT REPEAT 117 140 LRR 4.
FT REPEAT 141 164 LRR 5.
FT REPEAT 165 188 LRR 6.
FT REPEAT 224 247 LRR 7.
FT REPEAT 249 272 LRR 8.
FT REPEAT 273 296 LRR 9.
FT REPEAT 297 320 LRR 10.
FT REPEAT 321 343 LRR 11.
FT REPEAT 344 367 LRR 12.
FT REPEAT 391 391 LRR 13.
FT REPEAT 395 418 LRR 14.
FT REPEAT 419 444 LRR 15.
FT REPEAT 446 469 LRR 16.
FT REPEAT 491 513 LRR 17.
FT REPEAT 514 537 LRR 18.
FT REPEAT 539 562 LRR 19.
FT REPEAT 615 638 LRR 20.
FT REPEAT 640 663 LRR 21.
FT CARBOHYD 196 196 (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 262 (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 454 454 (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 488 488 (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 530 530 (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 618 618 (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 648 648 (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 667 667 (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 782 AA: 88561 MW: 31146082D4527917 CRC64:

```

```

Query Match 9.3%; Score 163; DB 1; Length 782;
Best Local Similarity 26.7%; Pred. No. 1.1e-05;
Matches 86; Conservative 37; Mismatches 115; Indels 84; Gaps 14;

```

```

Db 138 NLKSLKRLNKGKNTATVATETQNLPELD-----LDLAYN 174
OY 81 DDEALTGQIPPOVDLPYLOA-LMFRKLPNLF--GKIPREISALDKSLRSLSTSLGP 137
Db 175 SISLSDNITDQVSGSLMFVNMHNKLNLYVAPSVPE-----QDTGLG- 221
OY 138 VPLFFPOLTKLCLDLSFNKLLGY-----PRLSTLPNKLALHLENNELTGETIPDIFGN 192
Db 222 -----LQNIKYLDSLFSNNITSAKOFPRPVELS-----LMQLYLGHNKLNLNATKDLFGN 270
OY 193 FAGSPDIYLSHNOITGVPRKTFARADPIR-LDFSGNRLGEDISFLGPKRKLEMLDFSGN 251
Db 271 MPRLOVLDLSHNSLYELDFOTFRNTKQLQWLDTSHNRISLTPNDLFRFLGNLRIVDFSHN 330
OY 252 VLSNFSRVOEFPSPS-----LTYLDLNHNQI-----SGSLSS--ELAKLDLOTGVNSD 297
Db 331 -----RLNSLTPNLFRETGLERLDVSHNILGKLPRLSTLSLSAQTLSLSDLSMNSIS- 382
OY 298 NNLCGKIPTGSGNLOREPRITAVL 319
Db 383 -----SLSHGQGLARFKCLSLW 399

```

```

RESULT 14
GP46_LEIAM STANDARD; PRT; 476 AA.
ID GP46_LEIAM
AC P21978;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SURFACE MEMBRANE GLYCOPROTEIN GP46/M-2 PRECURSOR.
OS Leishmania amazonensis.
OC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5659;
RN [1]

```

```

RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-61.
RC STRAIN-LTB0016;
RX MEDLINE-91046003; PubMed-2236047;
RA Lohman K.L., Langer P.J., McMahon-pratt D.;
RT "Molecular cloning and characterization of the immunologically
protective surface glycoprotein GP46/M-2 of Leishmania amazonensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8393-8397(1990).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC
CC EMBL: M38368; AAA29234.1; -.
DR PIR: A36478; A36478.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR003592; LRR-Out.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00370; LRR; 3.
KW Antigen; Membrane; GPI-anchor; Signal; Glycoprotein; Repeat.
FT CHAIN 1 32 SURFACE MEMBRANE GLYCOPROTEIN GP46/M-2.
FT PROPEP 421 476 REMOVED IN MATURE FORM.
FT LIPID 420 420 GPI-ANCHOR (POTENTIAL).
FT DOMAIN 107 202 4 X 24 AA TANDEM REPEATS.
FT REPEAT 107 130 1.
FT REPEAT 131 154 2.
FT REPEAT 155 178 3.
FT REPEAT 179 202 4.
FT CARBOHYD 224 224 G -> E (IN REF. 1; AA SEQUENCE).
FT CONFLICT 40 40 G -> T (IN REF. 1; AA SEQUENCE).
FT CONFLICT 44 44 K -> T (IN REF. 1; AA SEQUENCE).
SQ SEQUENCE 476 AA: 50136 MW: 722F5DDC48C308BB CRC64:

```

```

Query Match 9.1%; Score 160.5; DB 1; Length 476;
Best Local Similarity 23.7%; Pred. No. 8.8e-06;
Matches 54; Conservative 37; Mismatches 92; Indels 45; Gaps 8;

```

```

OY 9 PLICTCMIFCLPNLASORCNNDKALLQIT-----AKNFTIDSWSDCCG 61
Db 13 PLAAVALLLCTSSAPVARAGTSDFTGAQKNLTLYLOAFARALPDLDTW-TGSDFC 71
OY 62 WDLVECDENRIILSLIDDEALTGQIPPOVDLPYLOALMFRKLPNLFKIPREISAL 121
Db 72 WEHITCYSSG-----VG-----VMMHNV-DYTGTLREMPASV 102
OY 122 --KDKSLRLS--STLSGVPVLPFPPOLTKLCLDLSFNKLLGYIIPOLSTLPNKLAL 175
Db 103 DYKDVAILALDFGACQGLGTLPSPSSMKHLIYLDLDTGKTSVGLPPEWSEMTSAEAL 162
OY 176 HLERNELTGEIPDIPGPNFAGSPDIYLSHNOITGVPRKTFARADPIRLD 223
Db 163 QLENGCLSGSLPTSSMPKRLIVSLGNHFCGCPDPSKREDD--RLD 208

```

```

RESULT 15
FMOD_MOUSE STANDARD; PRT; 376 AA.
ID FMOD_MOUSE
AC P50608;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE FIBROMODULIN PRECURSOR (FM).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2002, 20:56:50 ; Search time 56.08 Seconds  
(without alignments)  
865.949 Million cell updates/sec

Title: US-09-308-140-7  
Perfect score: 1756  
Sequence: 1 MNIESRCPILICMIFCL.....FDRATVHNSCLGAPLPEC 332

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: SP-archaea:\*
  - 2: SP-bacteria:\*
  - 3: SP-fungi:\*
  - 4: SP-human:\*
  - 5: SP-invertebrate:\*
  - 6: SP-mammal:\*
  - 7: SP-mhc:\*
  - 8: SP-organelle:\*
  - 9: SP-phage:\*
  - 10: SP-plant:\*
  - 11: SP-rodent:\*
  - 12: SP-virus:\*
  - 13: SP-vertebrate:\*
  - 14: SP-unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1756	100.0	332	10	092438 daucus caro
2	937.5	53.4	327	10	038695 actinidia c
3	894	50.9	333	10	09AXP4 Q9AXP4 vilis vinif
4	877.5	50.0	330	10	P93270 malus domes
5	866.5	49.3	330	10	09LKO8 Q9LKO8 prunus maha
6	851	48.5	327	10	022522 prunus arme
7	851	48.5	327	10	080399 citrus sp.
8	851	48.5	327	10	09SAR2 Q9SAR2 citrus sp.
9	850	48.4	327	10	080420 citrus unsh
10	849	48.3	327	10	096401 citrus sine
11	844	48.1	327	10	040160 lycopersico
12	843	48.0	327	10	064465 citrus jamb
13	842.5	48.0	327	10	09SAY9 Q9SAY9 citrus iyo
14	839.5	47.8	327	10	09ZWH6 Q9ZWH6 poncirus tr
15	838	47.7	327	10	09ZWH5 Q9ZWH5 fortunella
16	830.5	47.3	327	10	080421 citrus iyo
17	819	46.6	329	10	038738 antirrhinum
18	813	46.3	327	10	064460 citrus jamb
19	808	46.0	303	10	09SED5 Q9SED5 eucalyptus

## ALIGNMENTS

20	808	46.0	303	10	Q9SED2	Q9SED2 eucalyptus
21	807	46.0	303	10	Q9SED1	Q9SED1 eucalyptus
22	802	45.7	303	10	Q9SED3	Q9SED3 eucalyptus
23	794	45.2	303	10	Q9SED4	Q9SED4 eucalyptus
24	792.5	45.1	330	10	Q9MSY9	Q9MSY9 arabidopsis
25	784	44.6	326	10	Q9MSY8	Q9MSY8 arabidopsis
26	784	44.6	330	10	Q9FLS8	Q9FLS8 arabidopsis
27	719	40.9	324	10	Q9M7E7	Q9M7E7 arabidopsis
28	718.5	40.9	325	10	Q9LH52	Q9LH52 arabidopsis
29	699	39.8	253	10	Q9LUX1	Q9LUX1 prunus pyrif
30	630	35.9	313	10	Q9XHD6	Q9XHD6 glycine max
31	575	32.7	313	10	Q9J885	Q9J885 glycine max
32	457.5	26.1	214	10	Q9LH53	Q9LH53 arabidopsis
33	439	25.0	203	10	Q9C7D9	Q9C7D9 arabidopsis
34	431.5	24.6	372	10	Q9LHK0	Q9LHK0 arabidopsis
35	411.5	23.4	478	10	Q9FWA8	Q9FWA8 arabidopsis
36	408.5	23.3	371	10	Q9LXU5	Q9LXU5 arabidopsis
37	401	22.8	480	10	Q48705	Q48705 arabidopsis
38	398.5	22.7	365	10	Q9L739	Q9L739 arabidopsis
39	398.5	22.7	512	10	Q9FWA7	Q9FWA7 arabidopsis
40	384	21.9	589	10	Q9FHL8	Q9FHL8 arabidopsis
41	378	21.5	976	10	Q42371	Q42371 arabidopsis
42	370	21.1	486	10	Q9LQ21	Q9LQ21 arabidopsis
43	370	21.1	907	10	Q9FWA9	Q9FWA9 arabidopsis
44	370	21.1	943	10	Q9SR17	Q9SR17 arabidopsis
45	363.5	20.7	1036	10	Q9FN37	Q9FN37 arabidopsis

RESULT 1  
ID 082438 PRELIMINARY: PRT; 332 AA; 1756 bits  
AC 082438: 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE ANTI-FREEZE PROTEIN.  
GN AFP.  
OS Daucus carota (Carrot).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asterales; euasterids II; Apiales; Apiaceae; Daucus.  
OX NCBI\_TaxID=4039;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. AUTUMN KING; TISSUE=TAP ROOT;  
RX MEDLINE=98429644; PubMed=9756474;  
RA Worral D., Elias L., Ashford D., Smallwood M., Sidebottom C.,  
RT Lillford P., Telford J., Holt C., Bowles D.;  
RT "A carrot leucine-rich-repeat protein that inhibits ice  
recrystallization.";  
RL Science 282:115-117(1998).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NEWMARKET FL;  
RX MEDLINE=99229753; PubMed=1021940;  
RA Meyer K., Keil M., Naldrett M.;  
RT "A leucine-rich repeat protein of carrot that exhibits antifreeze  
activity.";  
RL FEBS Lett. 447:171-178(1999).  
DR EMBL: AF055480; AAC62932.1;  
DR EMBL: AJ131340; CAB37347.1;  
DR EMBL: 33027; Dauca/2327/33027.  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR003592; LRR\_out.  
DR Pfam: PF00560; LRR; 4.  
DR SMART: SM00370; LRR; 4.  
DR Signal.  
SQ SEQUENCE 332 AA; 36845 MW; EEEF6FAEBEAE202 CRC64;

Query Match 100.0%; Score 1756; DB 10; Length 332;  
 Best Local Similarity 100.0%; Pred. No. 9.6e-146;  
 Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNIESFCPIICICMIFLCPLNLSASQRCNNNDKQALLQITALKNPITDTSWVSDDDCC 60  
 DB 1 MNIESFCPIICICMIFLCPLNLSASQRCNNNDKQALLQITALKNPITDTSWVSDDDCC 60  
 QY 61 GMDLVECDTSNRITSLITODDEALTGQIPPOVGDLPYLQALMFRKLPNLFGKIPPEISA 120  
 DB 61 GMDLVECDTSNRITSLITODDEALTGQIPPOVGDLPYLQALMFRKLPNLFGKIPPEISA 120  
 QY 121 LKDKSLRSLSTSLSGVPPLFPOLTKLTCLDLSFNKLLGVIPOLSTLPNLKALHLERN 180  
 DB 121 LKDKSLRSLSTSLSGVPPLFPOLTKLTCLDLSFNKLLGVIPOLSTLPNLKALHLERN 180  
 QY 181 ELTGEIPDIFGNFAGSPDLYLSHNLGTGFVPTFARADPIRLDFSGNRLGDISFLFGPK 240  
 DB 181 ELTGEIPDIFGNFAGSPDLYLSHNLGTGFVPTFARADPIRLDFSGNRLGDISFLFGPK 240  
 QY 241 KRLEMLDSSGNVLSNFRVQEPFSLTYLDLNHQISGLSSELAKLDLOTFNVSNNL 300  
 DB 241 KRLEMLDSSGNVLSNFRVQEPFSLTYLDLNHQISGLSSELAKLDLOTFNVSNNL 300  
 QY 301 CGKIPTGSLQRFDRFATVLAHNSCLGAPLPEC 332  
 DB 301 CGKIPTGSLQRFDRFATVLAHNSCLGAPLPEC 332

RESULT 2  
 ID 038695 PRELIMINARY; PRT: 327 AA.  
 AC 038695;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE POLYGALACTURONASE INHIBITOR PRECURSOR.  
 OS Acetabularia chinensis (kiwi) (Yangtao).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; Ericales; Actinidiaceae; Actinidia.  
 OX NCBI\_TaxID=3625;  
 RN [1]  
 RP SEQUENCE FROM N. A.  
 RC STRAIN=HAYWARD; TISSUE=UNRIPE FRUIT;  
 RC Simpson C.G., Gardner R.C.;  
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z45053; CA88846.1;  
 DR MendeL: 14029; Accch: 2327; 14029.  
 DR InterPro: IPR001611; LRR.  
 DR Pfam: PF00560; LRR\_6.  
 DR SMART: SM00370; LRR\_6.  
 DR Signal.  
 FT CHAIN 1 21 POTENTIAL.  
 FT CDSIN 22 327 POLYGALACTURONASE INHIBITOR.  
 SQ SEQUENCE 327 AA; 36163 MW; 9335E02978378765 CRC64;

Query Match 53.4%; Score 937.5; DB 10; Length 327;  
 Best Local Similarity 59.9%; Pred. No. 4.3e-74;  
 Matches 191; Conservative 35; Mismatches 90; Indels 3; Gaps 3;

QY 15 MIFCLPLNLSASQRCNNNDKQALLQITALKNPITDTSWVSDDDCCGMDLVECDTSNR 74  
 DB 10 LRFSLSPSLSDSCNDRKDKVLLRIKQALNLPYLASWNPNDCCMYNVDCDLTNRI 69  
 QY 75 ISLIIDDEALTGQIPPOVGDLPYLQALMFRKLPNLFGKIPPEISALKDKLSLSTSL 134  
 DB 70 IALTIFSGN-ISCQIPAAVGDLPYLQTLIFRKLSTNLGQIPSAISKLSNKMVRLSWNL 128  
 QY 135 SGVPPLFPOLTKLTCLDLSFNKLLGVIPOLSTLPNLKALHLERNELTGEIPDIFGNA 194

DB 129 SGVPSPFSQKLNLTFLDLSFNDLTGSTIPSSSKLTNLDALHDKNKLGTGIPNSFGFT 188  
 QY 195 GS-PDIYLSHNLGTGFVPTFARADPIRLDFSGNRLGDISFLFGPKRLEMLDFSGNVL 253  
 DB 189 GQVDDLYLSHNLGTGIRKTLGDLNFTVIDVSNNLSDGISFMGSNNKTQIDVPSNRKF 248  
 QY 254 SFNFSRVOEFPFSLTYLDLNHQISGLSSELAKLDLOTFNVSNNLCKGKIPFGNLDGF 313  
 DB 249 QFDLSKV-VFPOSLSLTDLNHNKIYGSLPVGLTKLDLYLVNVSNNRCLGILPTGKLDGF 307  
 QY 314 DRFAYLHNSCLGAPLPEC 332  
 DB 308 DOTSFHNRCLGAPLPDC 326

RESULT 3  
 ID 09AXP4 PRELIMINARY; PRT: 333 AA.  
 AC 09AXP4;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE POLYGALACTURONASE INHIBITING PROTEIN.  
 OS Vitis vulpifera (Grape).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;  
 OC Vitis.  
 OX NCBI\_TaxID=29760;  
 RN [1]  
 RP SEQUENCE FROM N. A.  
 RC STRAIN=CV, CHARDONNAY; TISSUE=LEAVES;  
 RC Bezier A., Lambert B., Baillieu F.;  
 RT "Molecular cloning of a grapevine gene coding for polygalacturonase  
 inhibiting protein."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF305093; AK114075.1;  
 SQ SEQUENCE 333 AA; 36975 MW; 57BB840E1096C4FD CRC64;

Query Match 50.9%; Score 894; DB 10; Length 333;  
 Best Local Similarity 56.0%; Pred. No. 2.8e-70;  
 Matches 187; Conservative 40; Mismatches 101; Indels 6; Gaps 4;

QY 3 IESSFCPIICICMIFLCPLNLSASQRCNNNDKQALLQITALKNPITDTSWVSDDDCC 59  
 DB 1 METSKLFLSSLLVLLATRPCSLSRCPKDKVLLQIKKALDPYILASNNPNTDC 60  
 QY 60 GMDLVECDTSNRITSLITODDEALTGQIPPOVGDLPYLQALMFRKLPNLFGKIPPEIS 119  
 DB 61 CGWYCECDLTHRHNSLTJFSGQ-LSGOIPDAVGDDPFLFTLIFRKLSTNLGTGIPPAIA 119  
 QY 120 ALKDLKSLRSLSTSLSGVPPLFPOLTKLTCLDLSFNKLLGVIPOLSTLPNLKALHLER 179  
 DB 120 KLKHLKMWRLSWNLFGVPVAFSELNKLTLYLDSFNNSLSPIFGSLPLNKLALHIDR 179  
 QY 180 NELTGEIPDIFGNFAGS-PDIYLSHNLGTGFVPTFARADPIRLDFSGNRLGDISFLFG 238  
 DB 180 NHLGPIPDSEFGKAGSTPGLHLSHNLGSKIPYSFSGFDENVMWDLERNKLEGDLSIFN 239  
 QY 239 PKRLEMLDSSGNVLSNFRVQEPFSLTYLDLNHQISGLSSELAKLDLOTFNVSNN 298  
 DB 240 ANKSTQIVDSERNLFOFDLSRV-EFPKSLTSLDSHNKIGSLPEMWTSLDLOFLVNSYN 298  
 QY 299 NLGKIPTGSLQRFDRFATVLAHNSCLGAPLPEC 332  
 DB 299 RLCGKIPVGGKLSFDYDSYFHNRCCLGAPLQSC 332

RESULT 4  
 ID P93270 PRELIMINARY; PRT: 330 AA.  
 AC P93270;  
 ID P93270;



QY 63 DLVECDSTSNRIISLIQDDEALTGOIPPOVGDLPYLQALWFRKLPNLFKGIPEEISALK 122  
 Db 61 YCVTGDSTNRINSITIFAGQ-VSGQIPPOVGDLPYLETFEHNKQPNLTGPIQPSIAKLK 119  
 QY 123 DKSRLSTSGVPVLPFPOLTKLTDLSFNKLGVIPQSLTPNLKALHNERNEL 182  
 Db 120 LKELRLSWTNISGVPDPLSOLKNLTFDLSFNSMLTGSIPSWLQSLPNNLALRVDRNKL 179  
 QY 183 TGEIPDIFGNFAGS-PDIYLSHNLQTLGVFPKTFARADPIRLDFSGNRLEGGISLFGPKK 241  
 Db 180 TGHIPKSFGEFGDGVDPDLTSLHNLQSLGTPTSLAKLNTSIDFSNKLGGASMTFGINK 239  
 QY 242 RLEMDFSGNVLSENFSAVOEPPSLTYLDLNHNOISGSLSELAKLDLQTFNVSNNMLC 301  
 Db 240 TQQLVDSLNRNLLEILNSLV-EFSKSLTSLDLNHNKLTGIGVGLQVLDQFLNLSYNMLC 298  
 QY 302 GKIPGNGLRFRDRAVYAHNSCLGAPLPEC 332  
 Db 299 GQIPVGKLGQSFDSSTYFHNRCILGAPLPSC 329

RESULT 7  
 ID 080399 PRELIMINARY; PRT; 327 AA.  
 AC 080399;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE POLYGALACTURONASE-INHIBITING PROTEIN.  
 GN PGIP A.  
 OS Citrus sp. cv. sannumphung.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Sapindales; Rutaceae; Citrus.  
 OX NCBI\_TaxID=77791;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. SANNUMPHUNG;  
 RA Nalumpang S., Akimitsu K., Yamamoto H.;  
 RT "Polygalacturonase-inhibiting protein (PGIP) gene of citrus (cv. sannumphung)".  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB015356; BAA29024.1;  
 DR Mendel: 31062; Clisp:2327;31062.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003592; LRR\_out.  
 DR Pfam: PF00560; LRR; 4.  
 DR SMART: SM00370; LRR; 5.  
 SO SEQUENCE 327 AA; 36177 MW; E40BEDD6C0ED87A5 CRC64;

Query Match 48.5%; Score 851; DB 10; Length 327;  
 Best Local Similarity 55.1%; Pred. No. 1.6e-66;  
 Matches 177; Conservative 38; Mismatches 96; Indels 10; Gaps 5;  
 QY 11 LCICMFLCLPNLSASORCNNDKQALQIKTALKNPTTDSWSDCCGMDLVECDT 70  
 Db 13 LCICL-----SPSLSDLCNPDKVKVLLFKKSLNNPYVLASWNPKTCCDWCYVCULT 66  
 QY 71 SNRIISLIQDDEALTGOIPPOVGDLPYLQALWFRKLPNLFKGIPEEISALKDLSELRIS 130  
 Db 67 TNRINSITIFAGD-LPGQIPPEVGDLPYLETFEHNKQPNLTGPIQPSIAKLK 125  
 QY 131 STLSGVPVLPFPOLTKLTDLSFNKLGVIPQSLTPNLKALHNERNELTGEIPDI 190  
 Db 126 WTNISGVPVDFIRQTLNLTLELSEFNNSLGTIPGSLKLOKLGALHLDNRNKLTSIPSEF 185  
 QY 191 GNFAGS-PDIYLSHNLQTLGVFPKTFARADPIRLDFSGNRLEGGISLFGPKKRLMLDFS 249  
 Db 186 GTFGSIPDLTSLHNLQSLGIPASLGSMDNTIDLSRNKLEGGASLFGLNKKTORIDVS 245  
 QY 250 GNVLSFNFSRVOEPPSLTYLDLNHNOISGSLSELAKLD-LQTFNVSNNMLCGKIPG 308  
 Db 250 GNVLSFNFSRVOEPPSLTYLDLNHNOISGSLSELAKLD-LQTFNVSNNMLCGKIPG 308

Db 246 RNLEFNLSKV-EFPQSLTNLDLNHNKLTGSIQPSIAKLHNLGLNVSYNRLCGPIVWG 304  
 QY 309 NLQRPDRTAYLHNSCLGAPL 329  
 Db 305 KLGSGYTEYEFHNRCILGAPL 325

RESULT 8  
 ID 09SA22 PRELIMINARY; PRT; 327 AA.  
 AC 09SA22;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE POLYGALACTURONASE-INHIBITING PROTEIN.  
 GN PGIP B.  
 OS Citrus sp. cv. sannumphung.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Sapindales; Rutaceae; Citrus.  
 OX NCBI\_TaxID=77791;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. SANNUMPHUNG;  
 RA Nalumpang S., Akimitsu K., Yamamoto H.;  
 RT "Polygalacturonase-inhibiting protein (PGIP B) gene of citrus (cv. sannumphung)".  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB015643; BAA29056.1;  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003592; LRR\_out.  
 DR Pfam: PF00560; LRR; 4.  
 DR PRINTS: PR00019; LEURICHRPT.  
 DR SMART: SM00370; LRR; 5.  
 SO SEQUENCE 327 AA; 36168 MW; 86BA2C84C48A7177 CRC64;

Query Match 48.5%; Score 851; DB 10; Length 327;  
 Best Local Similarity 55.1%; Pred. No. 1.6e-66;  
 Matches 177; Conservative 38; Mismatches 96; Indels 10; Gaps 5;  
 QY 11 LCICMFLCLPNLSASORCNNDKQALQIKTALKNPTTDSWSDCCGMDLVECDT 70  
 Db 13 LCICL-----SPSLSDLCNPDKVKVLLFKKSLNNPYVLASWNPKTCCDWCYVCULT 66  
 QY 71 SNRIISLIQDDEALTGOIPPOVGDLPYLQALWFRKLPNLFKGIPEEISALKDLSELRIS 130  
 Db 67 TNRINSITIFAGD-LPGQIPPEVGDLPYLETFEHNKQPNLTGPIQPSIAKLK 125  
 QY 131 STLSGVPVLPFPOLTKLTDLSFNKLGVIPQSLTPNLKALHNERNELTGEIPDI 190  
 Db 126 WTNISGVPVDFIRQTLNLTLELSEFNNSLGTIPGSLKLOKLGALHLDNRNKLTSIPSEF 185  
 QY 191 GNFAGS-PDIYLSHNLQTLGVFPKTFARADPIRLDFSGNRLEGGISLFGPKKRLMLDFS 249  
 Db 186 GTFGSIPDLTSLHNLQSLGIPASLGSMDNTIDLSRNKLEGGASLFGLNKKTORIDVS 245  
 QY 250 GNVLSFNFSRVOEPPSLTYLDLNHNOISGSLSELAKLD-LQTFNVSNNMLCGKIPG 308  
 Db 246 RNLEFNLSKV-EFPQSLTNLDLNHNKLTGSIQPSIAKLHNLGLNVSYNRLCGPIVWG 304  
 QY 309 NLQRPDRTAYLHNSCLGAPL 329  
 Db 305 KLGSGYTEYEFHNRCILGAPL 325  
 RESULT 9  
 ID 080420 PRELIMINARY; PRT; 327 AA.  
 AC 080420;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

DE POLYGLACTURONASE INHIBITOR (PGIP).  
 GN Citrus unshiu (Satsuma orange).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Sapindales; Rutaceae; Citrus.  
 NCBI\_TaxID=55188;  
 RX  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Golch Y., Akimitsu K., Yamamoto H.;  
 RT "Citrus unshiu PGIP gene for polygalacturonase inhibitor."  
 RL Submitted (Jul-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB016204; AAA31841.1;  
 DR Mendel: 31083; Citrus\_2327;31083.  
 DR InterPro: IPR001611; LRR.  
 DR Pfam: PF00560; LRR; 4.  
 DR SMART: SM00370; LRR; 5.  
 SQ SEQUENCE 327 AA; 36147 MW; E15A8C7C90A19D5 CRC64;

Query Match 48.4%; Score 850; DB 10; Length 327;  
 Best Local Similarity 55.1%; Pred. No. 1.9e-66;  
 Matches 177; Conservative 38; Mismatches 96; Indels 10; Gaps 5;

OY 11 LCICMIFCLPMLASORCNRNNNDKQALLQIKTALKNPTTDSWSDCCGMDLYECDET 70  
 DB 13 LCLCT-----SPSLSDICNPDKRVLLFKFKSLNPPVLA5MNPKTDCDDWYCYTCDLT 66  
 OY 71 SNRIISLIIDDEALTGQIPPOVGDPYLQALMFKRLPNLFGKIPPEISALDKLSRLS 130  
 DB 67 TNRIISLTIFACD-LPGQIPPEVGDPYLETLMFKRLPSLTGPIQALAKLKNLTLRIS 125  
 OY 131 STLSGVPVLPFPQTLTKLCTLDLSFNKLGVIPPOLSTLPNLKALHLERNELTGEIPDIF 190  
 DB 126 WNIISGVPVDFIROLTNLTFLLESLFNLSGTLPGSLSKLOKLGALHLDRNKLTSIPESF 185  
 OY 191 GNFAGS-PDIYLSHNOLGTFVPKTFARADPIRLDPSGNLEEDISFLFPPKRLKLEMDFS 249  
 DB 186 GFTGTSIPDLYSHNOLSGKIPASLGSMDENTIDLSRNKLEGDASFGLNKTTRIDVS 245  
 OY 250 GNVLSFNFRVOEPFPLSLTYLDLNHNQISLSSELAELD-LQTFNVSNNLCGKIPITGG 308  
 DB 246 RULLEFNLSKV-EFPOSLTNLDLNHNKIFGSIQAOTTSLENLGFNVSYNRCLGPIPVGG 304  
 OY 309 NLQRFRTAYLHNSCLCGAPL 329  
 DB 305 KLSFGYTEYEFHNRCLCGAPL 325

RESULT 10  
 OY 096401 PRELIMINARY; PRT: 327 AA.  
 AC 096401:  
 DT 01-FEB-1997 (TREMblrel. 02, Created)  
 DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE POLYGLACTURONASE-INHIBITING PROTEIN.  
 OS Citrus sinensis (Sweet orange).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Sapindales; Rutaceae; Citrus.  
 NCBI\_TaxID=2711;  
 RX  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. HAWLIN; TISSUE=CALLUS;  
 RA Mayer R.T.;  
 RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Y08618; CAA69910.1;  
 DR Mendel: 14033; Cits1;2327;14033.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003592; LRR\_out.  
 DR Pfam: PF00560; LRR; 4.  
 SQ SEQUENCE 327 AA; 36147 MW; E15A8C7C90A19D5 CRC64;

DR SMART: SM00370; LRR; 5.  
 SQ SEQUENCE 327 AA; 36216 MW; 151F5C77C8BBA8D5 CRC64;

Query Match 48.3%; Score 849; DB 10; Length 327;  
 Best Local Similarity 55.1%; Pred. No. 2.4e-66;  
 Matches 177; Conservative 38; Mismatches 96; Indels 10; Gaps 5;

OY 11 LCICMIFCLPMLASORCNRNNNDKQALLQIKTALKNPTTDSWSDCCGMDLYECDET 70  
 DB 13 LCLCT-----SPSLSDICNPDKRVLLFKFKSLNPPVLA5MNPKTDCDDWYCYTCDLT 66  
 OY 71 SNRIISLIIDDEALTGQIPPOVGDPYLQALMFKRLPNLFGKIPPEISALDKLSRLS 130  
 DB 67 TNRIISLTIFACD-LPGQIPPEVGDPYLETLMFKRLPSLTGPIQALAKLKNLTLRIS 125  
 OY 131 STLSGVPVLPFPQTLTKLCTLDLSFNKLGVIPPOLSTLPNLKALHLERNELTGEIPDIF 190  
 DB 126 WNIISGVPVDFIROLTNLTFLLESLFNLSGTLPGSLSKLOKLGALHLDRNKLTSIPESF 185  
 OY 191 GNFAGS-PDIYLSHNOLGTFVPKTFARADPIRLDPSGNLEEDISFLFPPKRLKLEMDFS 249  
 DB 186 GFTGTSIPDLYSHNOLSGKIPASLGSMDENTIDLSRNKLEGDASFGLNKTTRIDVS 245  
 OY 250 GNVLSFNFRVOEPFPLSLTYLDLNHNQISLSSELAELD-LQTFNVSNNLCGKIPITGG 308  
 DB 246 RULLEFNLSKV-EFPOSLTNLDLNHNKIFGSIQAOTTSLENLGFNVSYNRCLGPIPVGG 304  
 OY 309 NLQRFRTAYLHNSCLCGAPL 329  
 DB 305 KLSFGYTEYEFHNRCLCGAPL 325

RESULT 11  
 OY 040160 PRELIMINARY; PRT: 327 AA.  
 AC 040160:  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE POLYGLACTURONASE INHIBITOR PROTEIN PRECURSOR.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 NCBI\_TaxID=4081;  
 RX  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VFWT CHERRY;  
 RA MEDLINE=94339475; PubMed=8061315;  
 RA Stoltz H.U., Contos J.J., Powell A.L., Bennett A.B., Labavitch J.M.;  
 RT "Structure and expression of an inhibitor of fungal polygalacturonases  
 from tomato."  
 RL Plant Mol. Biol. 25:607-617(1994).  
 DR EMBL: L26529; AAA53547.1;  
 DR Mendel: 14032; Lycos;2327;14032.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003592; LRR\_out.  
 DR Pfam: PF00560; LRR; 5.  
 DR SMART: SM00370; LRR; 4.  
 KW Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT CHAIN 21 327 POLYGLACTURONASE INHIBITOR PROTEIN.  
 SQ SEQUENCE 327 AA; 36490 MW; DCD61BF647D03FAL CRC64;

Query Match 48.1%; Score 844; DB 10; Length 327;  
 Best Local Similarity 52.8%; Pred. No. 6.5e-66;  
 Matches 171; Conservative 46; Mismatches 103; Indels 4; Gaps 3;

OY 11 LCICMIFCLPMLASORCNRNNNDKQALLQIKTALKNPTTDSWSDCCGMDLYECDET 70  
 DB 5 LLLVYIFLCFASPSL5VRCNPKDKRVLLQIKKDLGNPFIHLASMDPNTDCCYWIYIKCBK 64

